NIH Public Access
Author Manuscript
Future Microbiol. Author manuscript; available in PMC 2015 August 01.

Published in final edited form as:
Future Microbiol. 2014 October ; 9: 1319–1327. doi:10.2217/fmb.14.91.

Structure and function of bacteriophage T4

Moh Lan Yap[1] and **Michael G Rossmann**[1,*]

[1]Department of Biological Sciences, Purdue University, 240 S. Martin Jischke Drive, West Lafayette, IN 47907-2032, USA

Abstract

Bacteriophage T4 is the most well-studied member of Myoviridae, the most complex family of tailed phages. T4 assembly is divided into three independent pathways: the head, the tail and the long tail fibers. The prolate head encapsidates a 172 kbp concatemeric dsDNA genome. The 925 Å-long tail is surrounded by the contractile sheath and ends with a hexagonal baseplate. Six long tail fibers are attached to the baseplate's periphery and are the host cell's recognition sensors. The sheath and the baseplate undergo large conformational changes during infection. X-ray crystallography and cryo-electron microscopy have provided structural information on protein–protein and protein–nucleic acid interactions that regulate conformational changes during assembly and infection of *Escherichia coli* cells.

Keywords

assembly; bacteriophage T4; baseplate; contractile tail; genome packaging; T4 infection

Bacteriophage T4 is one of the seven *Escherichia coli* phages (T1–T7, T for type), which, in 1944, were suggested by Delbruck and coworkers to be models for study by the phage community [1]. Of these phages, the 'T-even' phages (T2, T4 and T6) are similar in structure, antigenic response and genetics [2]. Many early discoveries in molecular genetics such as the genetic code [3], and the confirmation that DNA is the genetic substance [4] were based on the study of these T even phages. In 1959, Brenner *et al.* [5] were able to obtain images of these phages using electron microscopy (EM). This initiated extensive studies of the highly symmetrical heads, tails and baseplates using primarily EM [6–18]. The complete genome of T4 was sequenced in 2003 [19]. In the same year, a higher resolution cryo-electron microscopy (cryoEM) structure was determined of the 'dome-shaped' baseplate, present in the infectious virus [20]. In the following year, the postinfection, star-shaped structure of the baseplate [21] and the detailed structure of the prolate head [22] were published. In the past 15 years, many T4 proteins have been studied using x-ray

© 2014 Future Medicine Ltd

*Author for correspondence: Tel.: +1 765 494 4911; Fax: +1 765 496 1189; mr@purdue.edu.

Financial & competing interests disclosure
The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.
No writing assistance was utilized in the production of this manuscript.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

crystallography. These were then fitted into lower resolution cryoEM maps, providing a glimpse of the virus at different stages of its life cycle.

Bacteriophage T4 is classified as a member in the Myoviridae family of the *Caudovirales* order because it has a contractile tail. The head, the tail and the long tail fibers (LTFs) of T4 are assembled independently before they are joined together to produce a mature phage (Figure 1). The 168 kbp dsDNA genome of T4 contains 289 open reading frames [19]. The mature phage has a 1150 Å-long and 850 Å-wide prolate head, which encapsidates the genome. A 925 Å-long and 240 Å-diameter contractile tail is attached to a special portal vertex at one end of the head. A 270 Å-high and 520 Å-diameter hexagonal baseplate is attached to the distal end of the contractile tail. Six 1450 Å-long LTFs are attached to the periphery of the hexagonal baseplate. The LTFs are the sensors that can recognize receptor molecules on the host. There are six short tail fibers (STFs) folded beneath the baseplate that unfold upon host recognition. After unfolding these STFs bind irreversibly to the host cell, thereby increasing the efficiency of infection. The contractile tail improves the efficiency of infection by making it possible for the tail tube to penetrate the outer host cell membrane prior to the delivery of phage DNA into the host cell.

Head structure

• Head proteins

The 194 MDa mature head encapsidates 172 kbp dsDNA, representing 102% of the unique genome [24,25]. The head is first assembled as an empty capsid that is subsequently packaged with DNA by an ATP-dependent packaging machine. This machine binds to the same special pentameric vertex that is later occupied by the phage tail. The head is prolate, meaning that it has two icosahedral ends and a cylindrical mid-section. The geometrical organization (expressed as triangulation numbers) of the ends and mid-section are based on planar hexagonal grids used by Caspar and Klug [26] and Moody [27] to predict the quasi-equivalent environments of the individual protein components. The icosahedral caps have $T_{end} = 13$ *laevo* symmetry, while the cylindrical mid-section is characterized by the triangulation number $T_{mid} = 20$ (Figure 2) [22]. The capsid is composed of 930 post-translationally modified monomers, or 155 hexamers of the major protein, gene product 23 (gp23*, where the * signifies post-translational cleavage) [24,28]. In addition, gp24* forms pentamers at 11 of the 12 pentameric vertices [29,30]. Gp24* is homologous to gp23* with 21% of the residues being identical [31]. The presence of proteins, homologous to the major capsid protein, which form pentamers as opposed to hexamers is a frequent solution to the formation of the pentameric vertices in icosahedral viruses [32–35]. The 12th vertex is the special vertex that is occupied by the dodecameric assembly of the portal protein gp20 [36,37]. The portal protein has multiple roles. It initiates head assembly, genome packaging and serves as the genome gatekeeper to prevent leakage of the packaged DNA [38]. Two accessory proteins, Hoc (highly antigenic outer capsid protein) and Soc (small outer capsid protein) attach to the capsid surface (Figure 2E) [29]. The rod-shaped Soc binds between two gp23* hexamers, thus forming a continuous mesh surrounding the hexameric gp23* on the capsid [39]. Soc maintains the stability of the head under extreme environments [40]. Hoc is an elongated molecule protruding from the center of gp23* hexamers [41]. Its Ig-like

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

domains, exposed on the outer surface of the head, may provide survival advantages to the phage.

The structure of the vertex protein gp24* is similar to the HK97 phage capsid protein with 2.7 Å rmsd between 149 equivalenced Cα atoms (Figure 3) [31]. Furthermore, the major capsid protein gp23* and the vertex protein gp24* are similar in molecular weight and in sequence. Thus, the capsid proteins of T4, like that of most tailed bacteriophages, have the HK97-like capsid protein fold. However, only 17% of the structurally aligned residues between T4 gp24 and the HK97 capsid protein were identical. The structurally aligned residues belong to the peripheral (P) and axial (A) domains. The A domains of the HK97 major capsid protein subunits form the central part of hexameric and pentameric capsomers, whereas the P domains are located in the capsomers' periphery. In comparison with the HK97 capsid protein, T4 gp24 has an insertion (I) domain (residues 82–142) in place of the elongated (E) loop (Figure 3A).

• Head morphogenesis

There are several stages in the assembly of the head (Figure 4): prohead formation, prohead proteolysis, DNA packaging, expansion of the prolate head and binding of the Hoc and Soc accessory proteins. The T4 prohead starts to assemble with the formation of the membrane-bound initiation complex, which comprises the portal protein gp20 mediated by the chaperone protein gp40 [42–46]. The initiation complex first attaches to the inner side of the host cytoplasmic membrane followed by binding of the prohead core (scaffold) proteins gp21, gp22, gp67, gp68, the initiation proteins IPI, IPII, IPIII and gpalt [47–51]. Subsequently, the capsid proteins gp23 and gp24 start to form a shell around the core proteins to assemble into a prohead. Both the core and shell assemble concurrently. The major capsid protein gp23 requires two chaperone proteins, the host-encoded GroEL and the phage-encoded gp31, for proper folding [52–54]. A total of 930 copies of the gp23 and 55 copies of gp24 are needed for the assembly of the prohead [24,28–29]. The volume of the prohead is approximately 15–20% smaller than that of the mature head [55–57]. The gp21 is then self-cleaved and activated to become a protease (T4PPase) in the prohead [58–62]. The amino terminal peptides of gp23, gp24, IPI, IPII, IPIII and gpalt are cleaved, whereas gp21, gp22, gp67 and gp68 are extensively digested into small peptide fragments. All digested peptide fragments except for peptide II and IV of gp67 and gp22, respectively, are then expelled from the prohead [29,63–64]. The cleaved head is released from the membrane into the cytosol and the gp23* and probably gp24* proteins undergo a large conformational rearrangement resulting in the expansion of the prohead. During this process, the capsid's facets are flattened and the wall of the capsid becomes thinner. The processed prohead is then released from the membrane for DNA packaging.

• The genome-packaging machine

The genome-packaging machine consists of three components: the dodecameric portal protein gp20, the pentameric large terminase gp17 and the small terminase gp16 (Figure 5) [38,66–67]. The external shape of the dodecameric portal protein assembly has a cylindrically shaped structure with the wider end inside the capsid, whereas the narrower end protrudes out of the capsid creating an attachment platform for the packaging motor

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(Figure 5B) [68–70]. The gp20 may act as a valve to stop the packaged DNA from escaping the head during successive strokes of the packaging motor and also serves as a sensor when a headful of genome is packaged [71,72]. The atomic structures of the portal proteins of phages φ29, SPP1 and P22 have been determined [68–70]. These portal proteins have less than 20% sequence identity, but are similar in their overall shape and structure.

The large terminase gp17 consists of two functional domains, the N-terminal ATPase domain and the C-terminal nuclease domain, connected by a flexible linker (Figure 5A) [73–75]. The ATPase domain consists of the classic nucleotide binding fold [76], the functional motifs such as Walker A, Walker B and catalytic carboxylate found in most ATP-driven macromolecular motors. There are two DNA-binding grooves on opposite sides of the C-terminal nuclease domain. One of these sites functions to cleave the concatenated DNA molecule to create an end for initiating packaging and again after packaging has been completed. The other binding site is most probably used to bind the DNA during DNA translocation into the head. A flexible peptide linker between the N-terminal ATPase domain and the C-terminal nuclease domain is essential for DNA translocation [77]. Five copies of gp17 assemble into a packaging motor on the protruding ‘stalk’ of the dodecameric portal, thus creating a symmetry mismatch between the portal and the motor.

The small terminase, gp16, is involved in initiating genome packaging [78,79] and regulating the gp17 functions [80]. In phages such as lambda and SPP1, the small terminase binds to a specific sequence (cos and pac sites, respectively) and brings it to the large terminase for initial cleavage to start the packaging mechanism [81,82]. However, there are no unique cos and pac sequences in the T4 genome [83,84]. Thus, gp16 may bind weakly to nonspecific DNA sequences to initiate DNA packaging. The functional oligomeric state of T4 gp16 is uncertain as the crystal structures of the small terminases vary from eight to 12 among three different phages (Sf6, SF6 and 44RR) [85–87]. Nevertheless, the domain organization of the small terminases is conserved and consists of an N-terminal DNA-binding domain, a central oligomerization domain, and a C-terminal large terminase-binding domain. The central oligomerization domain forms a ring-shaped structure. The N-terminal domains fold into a helix-turn-helix structure located around the periphery of the ring, whereas the C-terminal domains form a crown over the end of the ring assembly. Crystal structures, biochemical analyses and mutational studies suggest that the DNA wraps around the small terminase assembly [85–88].

• DNA-packaging mechanism

Genomic DNA is prepared for packaging when the gp17/gp16 complex cleaves the concatenated DNA and associates the cleaved end with an empty head [89,90]. The terminase-DNA complex is then attached to the platform created by the protruding end of the dodecameric portal protein gp20 at the unique vertex of the prohead, where a pentameric packaging motor is being assembled. The gp16 initiates the DNA translocation by stimulating the ATPase activity of gp17 before releasing itself from the packaging complex [87]. The capsid expands during the DNA packaging. The expansion causes rotation (and probably partial refolding) of the gp23* subunits. The distance between centers of neighboring gp23* capsomers increases from approximately 117 Å in the prohead to

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

approximately 140 Å in the mature capsid [15–16,91–95]. The expansion results in approximately 50% increase of the capsid volume [96]. The irreversible expansion stabilizes the capsid and creates binding sites for Soc and Hoc [40,97].

Various mechanisms have been proposed for the DNA-packaging process. A mechanism involving rotation of the portal complex was first proposed by Hendrix in 1978 [98], but found to be unacceptable because the motor continued to operate even after attachment of the gp20 monomers to the static shell of the capsid through Hoc [99]. Sun *et al.* [73] proposed a mechanism that depended on structural knowledge and electrostatic attraction or repulsion (Figure 6). They found two alternative conformational states of the gp17 subunit, identified as 'relaxed' and 'tensed'. In the tensed state, the N- and C-terminal domains of gp17 are bound together by means of a series of salt bridges as seen in the crystal structure of the complete gp17 molecule. In the relaxed state, the N- and C-terminal domains are separated and the flexible linker is in an extended conformation as observed by cryoEM of proheads complexed with gp17. DNA packaging occurs when the relaxed state becomes tensed, thereby electrostatically attracting the C-terminal domain toward the N-terminal domain that is bound to the prohead. The DNA, bound to the C-terminal domain of gp17, is pushed two bases further into the head, which puts the next two bases into register with the neighboring gp17 molecule related by the fivefold symmetry, consistent with the symmetry of the DNA structure and the observation that there are five gp17 molecules bound to the prohead (Figure 5).

The motor generates forces as high as approximately 60 pN and translocation rates of up to approximately 2000 bp/s [100]. The 172 kb T4 genome can be packaged in approximately 5 min. The DNA is tightly packed in concentric layers (Figure 6F) [101]. When the head is filled by DNA, a signal, which might be the internal pressure produced by the packaged DNA, is presumably sensed by the portal protein (and transmitted to the gp17). The nuclease domain of gp17 then cleaves the DNA and the pentameric gp17 motor is released from the capsid. The mature head is completed by attachment of the neck proteins gp13, gp14, gp2 and gp4 to the portal vertex. The head is then ready to be joined to the tail and LTFs to produce a mature phage.

Tail structure

• Tail morphogenesis

The tail consists of the tail tube surrounded by a helical sheath. Both tube and sheath are attached to the dome-shaped baseplate at the end away from the head. The six STFs are folded beneath the baseplate. Assembly of the tail begins with the assemblies of the six wedges and a central hub (Figure 7A & B). These together form the baseplate that then nucleates the assembly of the tail tube around which the sheath assembles (Figure 7C).

The tail assembly pathway was established by means of a series of experiments using sucrose density gradient (SDG) centrifugation assays of mutant-infected bacterial lysates in combination with *in vitro* complementation assay of the SDG fractions [103–105]. The SDG fractions were further investigated using EM to observe the partially assembled intermediate baseplate structures. Phage mutants lacking a gene product were used to analyze how the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

precursor complex accumulated in each mutant-infected bacterium. All of the gene products that eventually make the assembled tail remained in solution until intermediate precursor complexes were formed. They thus concluded that the wedge assembly is a strictly ordered sequential event. These results were later confirmed by Plishker *et al.* [106], Ferguson and Coombs [107] and Yap *et al.* [108].

The wedge proteins gp11, gp10, gp7, gp8, gp6, gp53 and gp25 are assembled sequentially in this order to form a wedge (Figure 7A). Six wedges were then associated around the central hub to form a hexagonal baseplate (Figure 7B). The baseplate assembly is completed by the attachment of STFs and gp9 at its periphery and, of gp48 and gp54 to the top of the hub, thus creating a platform for the polymerization of the tail tube (gp19) and the tail sheath (gp18) (Figure 7C). The tail tube probably polymerizes along the elongated ruler protein, gp29, which controls the length of the tail [109]. However, the mechanism of the tail tube polymerization and the incorporation of gp29 into the tail tube remains unclear. At the end of the tail tube, gp3 binds to the last row of gp19, thus stabilizing the polymerized tail tube. The tail sheath polymerizes around the tail tube in parallel with the formation of the tail tube. The tail assembly is completed by the binding of the tail terminator, gp15 to gp3 and the last row of gp18.

- **The structure of the baseplate**

During the infection process, the baseplate undergoes a dramatic conformational change that relaxes the high-energy dome-shaped structure (270 Å-high and 520 Å-diameter) (Figure 8A) [20] to a low-energy star-shaped structure (120 Å-high and 610 Å-diameter) (Figure 8B) [21]. The transition between these structures has been studied by analyzing the structures of the component proteins. Crystal structures of nine baseplate proteins (gp11, gp10, gp8, gp6, gp25, gp9, gp12, gp5 and gp27) (Figure 9) have been determined completely or partially. These structures were fitted into the 3D cryoEM reconstructed baseplate for both the dome-shaped and star-shaped conformations. Because each of the known crystal structures could be fitted unambiguously into both conformations, there cannot be a significant conformational change of the constituent proteins at the level of the 17 Å resolution of the cryoEM maps. Thus, these proteins must be sliding over each other while the conformation of the baseplate is changing from the dome-shaped to the star-shaped during the infection process.

Proteins at the periphery of the baseplate (gp11, gp10, gp9) have threefold symmetry (Figures 8F, 9C, E & F). These proteins are the adaptors to connect the LTFs to the baseplate and transmit a signal to the baseplate on recognizing a host cell surface lipopolysaccharide molecule. The adaptor for the attachment of the LTF is gp9, which is an elongated trimer [112] that has three domains: the N-terminal domain (residues 1–60) forms an α-helical, triple coiled coil; the middle domain (residues 61–166) consists of a seven-stranded β-sandwich; the C-terminal domain (residues 167–288) is an eight-stranded, antiparallel β-barrel structure. The C-terminal domain of gp9 associates with a trimeric LTF, whereas the N-terminal domain associates with the gp7 in the baseplate (Figure 8E). The hinge between the N-terminal and the middle domain provides flexibility to the LTFs when searching for host cell receptors. This change in gp9 probably initiates the baseplate

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

conformational change that eventually results in the transfer of the T4 genome into the host. For this process to be successful at least three LTFs need to lock the axis of the phage at a roughly normal orientation to the *E. coli* surface.

The trimeric gp10 is the largest protein component of each wedge of the baseplate. Based on the cryoEM density, it folds into four domains (I, II, III and IV). When the crystal structure of the C-terminal region (residues 397–602) was determined (Figure 9E), it was found that this segment belonged to what had been assigned as domains I and II [113] and had the overall appearance of an arrow. Domain I (residues 521–590) forms the shaft, domain II (residues 406–520 and 591–602) forms the arrowhead, domain III interacts with gp11 and domain IV binds to the STFs (Figure 8E & F).

The trimeric gp11 has three domains (Figure 9F) [114]. The N-terminal domain (residues 12–64) is an α-helical, coiled coil structure, the middle or 'finger' domain (residues 80–188) has a seven-stranded, antiparallel, skewed β-roll and an α-helix and the C-terminal domain (residues 65–79 and 189–218) consists of β-structure with multiple intermonomeric contacts for trimerization of the protein. The three fingers of gp11 clamp domain III of gp10 (Figure 8E & F). The kink in each STF is attached to the N-terminal domain and to one of the fingers of gp11. Thus, gp11 is holding the kink of the STFs and thereby helping the STFs to fold beneath the baseplate.

The STF protein, gp12, assembles as a 340 Å-long, club-shaped trimeric molecule [115,116]. The crystal structure of a fragment of gp12 (residues 85–527) forms a triple-β-helix structure (Figure 9G). The interior of the gp12 β-helix is hydrophobic. Unlike most STFs of other phages, T4 gp12 does not possess any enzymatic activity. Before infection, the six STFs are folded underneath the baseplate (Figure 8C), but when the virus has found a potential host, they unfold by straightening out at a bent region half way along their lengths (Figure 8F).

The crystal structures of the dimeric proteins gp6 and gp8 are known (Figure 9A & B). These proteins are near the center of the baseplate. The monomer of gp8 forms two domains, domain I (residues 1–87 and 246–334) and domain II (residues 88–245) [117]. The dimer of gp8 resembles a four-legged structure. Six dimeric gp6 molecules form a ring, which serve as the interaction interface between the neighboring wedges to form the hexagonal baseplate. Thus, the gp6 molecules maintain the integrity of the baseplate during the conformational change. A crystal structure of a part of gp6 shows that the C-terminal region can be divided into three domains [118].

Some monomeric proteins probably help to cement the baseplate proteins together. In particular, gp7 binds together the proteins in a wedge (Figure 8D & E), whereas gp53 and gp25 bind to the intersection of the wedges.

- **The cell-puncturing device**

The cell-puncturing device terminates the distal end of the tail tube and is therefore surrounded by the dome of the infectious virus (Figure 8C & D). This device acts not only as a mechanical way of inserting the phage tail tube into the periplasmic space, but also

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

contains three lysozyme domains for digesting the peptidoglycan layer within the periplasmic space. The cell-puncturing device consists of three proteins, gp5, gp27 and gp5.4. Both gp5 and gp27 are trimers forming a threefold symmetric prism-shaped structure when viewed from the top and resembles a flash light when viewed from the side (Figure 9H) [111]. The gp27 consists of four domains. Domain 1 (residues 2–111) and domain 3 (residues 297–239 and 307–368) have homologous structures that form a pseudo-sixfold symmetric trimeric torus. This structure binds to the inside of the dome-shaped base plate and forms a link between the sixfold symmetric dome and threefold symmetric puncturing device.

The gp5 component of the cell-puncturing device consists of three domains (Figure 9H). The N-terminal region forms an oligosaccharide-binding domain (residues 1–129), which interacts with gp27 to form the cylindrical 'head' of the flash light structure. The middle domain (residues 174–339) has lysozyme activity and a structure closely similar to hen egg white lysozyme [119]. The function of this domain is to lyse the host's periplasmic cell wall during penetration of the bacterial cell wall barrier. Three middle domains, linked by two long peptide linkers, surround the upper part of the trimeric β-helix cylinder of the C-terminal domain (residues 389–575). The gp5 lysozyme domain has the conserved residues Glu184, Asp193, Thr199 at the active site that correspond to the 'T4 lysozyme' (Glu11, Asp20, Thr26), required for phage exit, as well as hen egg white lysozyme (Glu35, Asp52). The C-terminal triple β-helix domain of gp5 contains 11 VXGXXXXX repeats that form a cylindrical hollow tube with an internal diameter of 28 Å and a length of 110 Å. The gp5 molecules undergo a maturation cleavage between residues Ser351 and Ala 352 to activate the lysozyme [120]. Both cleavage parts (gp5* and gp5C) remain attached by the C-terminal linker. This process increases the enzymatic activity by a factor of 10 [121]. At the distal end of the prism, there is bound an additional, needle-shaped protein that is probably gp5.4. Based on its sequence homology to a protein in bacterial type VI secretion (T6SS) [122], this protein is a monomer, containing a Zn atom at its end, coordinated by three histidines and one cysteine. The T4 phage can therefore be likened to a drill, with gp5.4 at the drill head that bores a hole into the outer cell membrane.

- **Tail tube protein & sheath**

Although the tail tube appears to have a smooth surface in most EM reconstructions, at an elevated contour level, the cryoEM map of the tail tube can be segmented into individual subunits [21]. The tail tube protein, gp19, is a monomeric protein in solution that assembles into the 925 Å-long tail tube (Figure 8B). The tube has a 90 Å external and 40 Å internal diameter [21].

The sheath (Figure 10) around the tail tube consists of 23 hexameric rings of gp18 that polymerize into a 240 Å-wide and 925Å-long, high-energy tube consisting of six right-handed helices. Each ring is 40.6 Å-high (the pitch) and is rotated by 17.2° (the twist) in a right-handed manner relative to the previous ring (Figure 10D) [110]. This conformation of the sheath collapses into the low energy form when the virus infects a host cell. During infection, the baseplate changes its conformation from a dome shape to a star shape structure. The sheath releases its energy, causing it to contract to a 330 Å-wide and 420 Å-

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

long structure. The contracted tail sheath has six right-handed helices with a pitch of only 16.4 Å and a twist angle of 32.9° situated between radii of 60 and 165 Å [21]. Thus, the phage head rotates by 345.4° about the tail axis to push the sharp tip of the tail, in the manner of a drill, to penetrate the outer cell membrane.

The tail sheath protein, gp18 (659 residues) is a monomeric protein in solution. The crystal structure of a major fragment (residues 1–510) of gp18 has been determined (Figure 10A) [123]. The sheath protein has four domains: domain I (residues 98–188) is a six-stranded β-barrel and an α-helix, domain II (residues 88–97 and 189–345) is a two-layer β-sandwich, flanked by four small α-helices, domain III (residues 24–87 and 346–510) has a β-sheet with five parallel and one antiparallel β-strands plus six α-helices surrounding the β-sheet and the fourth domain (residues 511–659) whose structure has not yet been determined. The fourth domain is positioned on the interior of the tail sheath, facing the tail tube, whereas domain I is protruding outward from the tail with no intersubunit contact. The N-terminal domain I of the first hexameric ring interacts with the baseplate proteins, gp48 and gp54 to initiate the sheath polymerization.

The shape of the gp18 molecules is the same in both the preinfection extended conformation and the postinfection contracted conformation, suggesting that the contraction occurs by sliding of the individual gp18 domains over each other (Figure 10B).

• The tail tube & tail sheath terminators

The polymerized tail tube and sheath are capped by the terminator proteins gp3 and gp15, respectively, to prevent depolymerization before the tail attaches to the head. Both gp3 and gp15 form hexameric rings that interact with the last row of gp19 and gp18 molecules. The crystal structure of gp15 [124], but not gp3 has been determined (Figure 10A). Gp15 has a four-stranded antiparallel β-sheet and an α-helix facing the inner side of the ring. The central pore and the side surface of gp15 are negatively charged, whereas the top and the bottom surfaces are positively charged. The top and bottom surfaces interact with gp14 and gp3 proteins, respectively. The interaction between gp15 and gp18 is different in the extended and contracted (postinfection) conformations (Figure 10B). In the contracted tail, the negatively charged side surface of the gp15 hexamer interacts with positively charged surfaces of the C-terminal domains of the gp18 molecules. These interactions help to maintain the integrity of the tail in its contracted form. The gp15 hexamer may have undergone a conformational change during the infection process, which might be propagated through gp14 and gp13, to the portal assembly to allow the release of the genomic DNA.

Long tail fibers collar & whiskers

The 1450 Å-long LTFs consist of four different proteins, namely gp34, gp35, gp36 and gp37 (Figure 11A & B) [13]. The chaperon protein gp57A is required for the trimerization of gp34 and gp37, whereas the chaperon protein gp38 is required for proper folding of gp37 [125]. The proximal half of the fiber is formed by gp34, which interacts with the adaptor protein gp9 on the baseplate [112]. The monomeric gp35 forms the hinge or knee connecting the proximal and distal parts of the LTFs. The proximal and distal half-fibers assemble independently (Figure 1). Subsequently, the C-terminal part of gp36 binds to the N-terminal

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

region of gp37 [126]. The distal part of the fiber is a trimer that can be divided by visual inspection of EM images into 11 domains (D1–D11) (Figure 11B) [127]. Domains D1 and D2 are a part of gp36 and D3–D11 are a part of gp37. The crystal structure of what had been assigned by EM as domains D10 and D11 of gp37 (residues 811–1026) has been determined (Figure 11C) [128]. This structure suggested that a more accurate description of this part of gp37 was in terms of a collar, needle and head domain. The head domain sits at the tip of the distal end and thus should be the T4 component that recognizes the receptor-binding site on the host cell [129]. Since gp37 is known to bind to lipopolysaccharide, and as protein saccharide interactions usually involve aromatic side chains, a series of Tyr and Trp residues at the tip of gp37 might be important for host recognition. Sequence analysis shows that the collar and needle domains are conserved among other phages, whereas the head domain has diverged, suggesting that the host range specificity is determined by the head domain, consistent with the head domain having the receptor recognition function.

The filled T4 head is attached to the tail via the neck proteins gp13 and gp14. Subsequently, six 500 Å-long, trimeric ‘whisker’ fibers (gpWac) are attached to the neck [130]. The whiskers have a trimeric coiled coil structure known as ‘fibritin’ (Figure 12A) [131,132]. The 50 amino acid residues at the N-terminus of gpWac bind to the neck region, whereas the 30 amino acid residues of the C-terminus are a chaperon that determines the proper trimerization of the fibritin [131–135]. Six other fibers, also constructed of fibritin, form a ‘collar’ around the neck (Figure 12B & C) [130]. The whisker and collar fibers are attached alternatively around the neck [124]. When the LTFs are retracted, the elbow regions of the fibers interact with the whiskers and the distal part of the fibers interacts with the collar.

Fibritin and gp63 (RNA ligase) promote attachment of the LTFs to the phage. Mutational studies showed that fibritin is also important in sensing the optimum environment for phage infection. In unfavorable conditions, the LTFs would be retracted and interact with the whiskers and collar, causing the proximal half of the LTFs to lie along the sheath and the distal half of the LTFs to angle out along the base of the head, preventing infection [136].

Conclusion

The structure of bacteriophage T4 has been extensively studied using various techniques such as complementation assays, cross-linking analyses, x-ray crystallography and cryoEM, providing a structural model of phage T4 at near atomic resolution. Structural similarities among phages as well as of T4 components with bacterial proteins, demonstrate common evolutionary ancestry or co-evolution with bacterial hosts. For example, the ATP-binding fold in the gp17 large terminase [137] is found in numerous situations that require an ATP-driven motor [76], or the HK97 fold [138] forms the basic unit to construct virus capsids of different sizes found among most tailed phages [31,139] and herpes viruses [140]. The T4 tail [111] is also found to resemble the bacterial type-VI secretion system [122] in its structural organization and function. Thus, the study of T4 and bacteriophages in general has provided basic biological and mechanistic information because of the generally highly conserved structural and functional themes over long evolutionary time spans.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Future perspective

The basic knowledge derived from studying the structure and assembly of T4 as well as the molecular mechanisms of its components could serve as the basis for adapting T4 toward medical applications such as an antibiotic agent, gene therapy, vaccine development and targeting of malfunctional cells. For instance, the T4 head could be engineered as a platform for displaying proteins to act as vaccines for variety of pathogens [141–146]. The accumulated and emerging knowledge could open the possibility to manipulate T4 for numerous medical applications.

Acknowledgements

The authors would like to thank Andrei Fokine and Venigalla B Rao for helpful comments and Sheryl Kelly for help in preparation of this manuscript.

The authors acknowledge support provided by NIH grant R01 AI081726 to MG Rossmann.

References

1. Anderson, TF. Electron microscopy of phages. In: Cairns, J.; Stent, GS.; Watson, JD., editors. Phage and the Origins of Molecular Biology. NY, USA: Cold Spring Harbor Laboratory Press; 1992. p. 63-78.
2. Demerec M, Fano U. Bacteriophage-resistant mutants in *Escherichia coli*. Genetics. 1945; 30(2): 119–136. [PubMed: 17247150]
3. Crick FH, Barnett L, Brenner S, Watts-Tobin RJ. General nature of the genetic code for proteins. Nature. 1961; 192(4809):1227–1232. [PubMed: 13882203]
4. Hershey AD, Chase M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. J. Gen. Physiol. 1952; 36(1):39–56. [PubMed: 12981234]
5. Brenner S, Streisinger G, Horne RW, et al. Structural components of bacteriophage. J. Mol. Biol. 1959; 1(3):281–292.
6. Bradley DE. The structure of coliphages. J. Gen. Microbiol. 1963; 31:435–445. [PubMed: 14014825]
7. DeRosier DJ, Moore PB. Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry. J. Mol. Biol. 1970; 52(2):355–369. [PubMed: 5485914]
8. Daems WT, van de Pol JH, Cohen JA. Some remarks on the morphology of bacteriophage T4B. J. Mol. Biol. 1961; 3(2):225–227. [PubMed: 13719284]
9. Fernandez-Moran, H. New approaches in the study of biological ultrastructure by high resolution electron microscopy. In: Harris, RJC., editor. The Interpretation of Ultrastructure. NY, USA: Academic Press; 1962. p. 411-427.
10. Kellenberger E, de la Tour EB. On the fine structure of normal and ‘polymerized’ tail sheath of phage T4. J. Ultrastruct. Res. 1964; 11:545–563. [PubMed: 14244176]
11. Krimm S, Anderson TF. Structure of normal and contracted tail sheaths of T4 bacteriophage. J. Mol. Biol. 1967; 27(2):197–202. [PubMed: 4167417]
12. Amos LA, Klug A. Three-dimensional image reconstruction of the contractile tail of T4 bacteriophage. J. Mol. Biol. 1975; 99(1):51–64. [PubMed: 1206701]
13. Epstein RH, Bolle A, Steinberg C, et al. Physiological studies of conditional lethal mutants of bacteriophage T4D. Cold Spring Harbor Symp. Quant. Biol. 1964; 28:375–392.
14. Finch JT, Klug A, Stretton AOW. The structure of the ‘polyheads’ of T4 bacteriophage. J. Mol. Biol. 1964; 10(3):570–575. [PubMed: 14257705]
15. DeRosier DJ, Klug A. Structure of the tubular variants of the head of bacteriophage T4 (polyheads). I. Arrangement of subunits in some classes of polyheads. J. Mol. Biol. 1972; 65(3): 469–488. [PubMed: 5023668]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

16. Yanagida M, DeRosier DJ, Klug A. Structure of the tubular variants of the head of bacteriophage T4 (polyheads). II. Structural transition from a hexamer to a 6+1 morphological unit. J. Mol. Biol. 1972; 65(3):489–499. [PubMed: 5023669]
17. Klug A, De Rosier DJ. Optical filtering of electron micrographs: reconstruction of one-sided images. Nature. 1966; 212(5057):29–32. [PubMed: 4164921]
18. Crowther RA, Lenk EV, Kikuchi Y, King J. Molecular reorganization in the hexagon to star transition of the baseplate of bacteriophage T4. J. Mol. Biol. 1977; 116(3):489–523. [PubMed: 592390]
19. Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, Ruger W. Bacteriophage T4 genome. Microbiol. Mol. Biol. Rev. 2003; 67(1):86–156. [PubMed: 12626685]
20. Kostyuchenko VA, Leiman PG, Chipman PR, et al. Three-dimensional structure of bacteriophage T4 baseplate. Nat. Struct. Biol. 2003; 10(9):688–693. [PubMed: 12923574]
21. Leiman PG, Chipman PR, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG. Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. Cell. 2004; 118(4):419–429. [PubMed: 15315755]
22. Fokine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG. Molecular architecture of the prolate head of bacteriophage T4. Proc. Natl Acad. Sci. USA. 2004; 101(16): 6003–6008. [PubMed: 15071181]
23. Leiman PG, Kanamaru S, Mesyanzhinov VV, Arisaka F, Rossmann MG. Structure and morphogenesis of bacteriophage T4. Cell Mol. Life Sci. 2003; 60(11):2356–2370. [PubMed: 14625682]
24. Baschong W, Baschong-Prescianotto C, Engel A, et al. Mass analysis of bacteriophage T4 proheads and mature heads by scanning transmission electron microscopy and hydrodynamic measurements. J. Struct. Biol. 1991; 106(2):93–101. [PubMed: 1804277]
25. Kim JS, Davidson N. Electron microscope heteroduplex study of sequence relations of T2, T4, and T6 bacteriophage DNAs. Virology. 1974; 57(1):93–111. [PubMed: 4594471]
26. Caspar DLD, Klug A. Physical principles in the construction of regular viruses. Cold Spring Harb. Symp. Quant. Biol. 1962; 27:1–24. [PubMed: 14019094]
27. Moody MF. Geometry of phage head construction. J. Mol. Biol. 1999; 293(2):401–433. [PubMed: 10529353]
28. Baschong W, Aebi U, Baschong-Preschianotto C, et al. Head structure of bacteriophages T2 and T4. J. Ultrastruct. Mol. Struct. Res. 1988; 99(3):189–202. [PubMed: 3198952]
29. Onorato L, Stirmer B, Showe MK. Isolation and characterization of bacteriophage T4 mutant preheads. J. Virol. 1978; 27(2):409–426. [PubMed: 691116]
30. Yanagida M, Suzuki Y, Toda T. Molecular organization of the head of bacteriophage T-even: underlying design principles. Adv. Biophys. 1984; 17:97–146. [PubMed: 6399817]
31. Fokine A, Leiman PG, Shneider MM, et al. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. Proc. Natl Acad. Sci. USA. 2005; 102(20):7163–7168. [PubMed: 15878991]
32. Burnett RM. The structure of the adenovirus capsid. II. The packing symmetry of hexon and its implications for viral architecture. J. Mol. Biol. 1985; 185(1):125–143. [PubMed: 4046035]
33. Stewart PL, Burnett RM, Cyrklaff M, Fuller SD. Image reconstruction reveals the complex molecular organization of adenovirus. Cell. 1991; 67(1):145–154. [PubMed: 1913814]
34. Rydman PS, Caldentey J, Butcher SJ, Fuller SD, Rutten T, Bamford DH. Bacteriophage PRD1 contains a labile receptor-binding structure at each vertex. J. Mol. Biol. 1999; 291(3):575–587. [PubMed: 10448038]
35. Yan X, Olson NH, Van Etten JL, Bergoin M, Rossmann MG, Baker TS. Structure and assembly of large lipid-containing dsDNA viruses. Nat. Struct. Biol. 2000; 7(2):101–103. [PubMed: 10655609]
36. Driedonks RA, Engel A, tenHeggeler B, van Driel R. Gene 20 product of bacteriophage T4. Its purification and structure. J. Mol. Biol. 1981; 152(4):641–662. [PubMed: 7334518]
37. Driedonks RA, Caldentey J. Gene 20 product of bacteriophage T4. II. Its structural organization in prehead and bacteriophage. J. Mol. Biol. 1983; 166(3):341–360. [PubMed: 6406677]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

38. Hsiao CL, Black LW. DNA packaging and pathway of bacteriophage T4 head assembly. Proc. Natl Acad. Sci. USA. 1977; 74(9):3652–3656. [PubMed: 269421]
39. Qin L, Fokine A, O'Donnell E, Rao VB, Rossmann MG. Structure of the small outer capsid protein, Soc: a clamp for stabilizing capsids of T4-like phages. J. Mol. Biol. 2010; 395(4):728–741. [PubMed: 19835886]
40. Ishii T, Yanagida M. The two dispensable structural proteins (soc and hoc) of the T4 phage capsid: their properties, isolation and characterization of defective mutants, and their binding with the defective heads *in vitro*. J. Mol. Biol. 1977; 109(4):487–514. [PubMed: 15127]
41. Fokine A, Islam MZ, Zhang Z, Bowman VD, Rao VB, Rossmann MG. Structure of the three N-terminal immunoglobulin domains of the highly immunogenic outer capsid protein from a T4-like bacteriophage. J. Virol. 2011; 85(16):8141–8148. [PubMed: 21632759]
42. Hsiao CL, Black LW. Head morphogenesis of bacteriophage T4. III. The role of gene 20 in DNA packaging. Virology. 1978; 91(1):26–38. [PubMed: 726265]
43. Hsiao CL, Black LW. Head morphogenesis of bacteriophage T4. II. The role of gene 40 in initiating prehead assembly. Virology. 1978; 91(1):15–25. [PubMed: 726259]
44. Hsiao CL, Black LW. Head morphogenesis of bacteriophage T4. I. Isolation and characterization of gene 40 mutants. Virology. 1978; 91(1):1–14. [PubMed: 726257]
45. Michaud G, Zachary A, Rao VB, Black LW. Membrane-associated assembly of a phage T4 DNA entrance vertex structure studied with expression vectors. J. Mol. Biol. 1989; 209(4):667–681. [PubMed: 2685327]
46. Black, LW.; Showe, MK.; Steven, AC. Morphogenesis of the T4 head. In: Karam, JD., editor. Molecular Biology of Bacteriophage T4. Washington, DC, USA: American Society for Microbiology; 1994. p. 218-258.
47. Van Driel R, Couture E. Assembly of bacteriophage T4 head-related strucutres. II. *In vitro* assembly of prehead-like structures. J. Mol. Biol. 1978; 123(2):115–128. [PubMed: 682197]
48. Van Driel R, Couture E. Assembly of the scaffolding core of bacteriophage T4 preheads. J. Mol. Biol. 1978; 123(4):713–719. [PubMed: 691063]
49. Caldentey J, Lepault J, Kellenberger E. Isolation and reassembly of bacteriophage T4 core proteins. J. Mol. Biol. 1987; 195(3):637–647. [PubMed: 3498844]
50. Traub, F.; Maeder, M.; Kellenberger, E. Bacteriophage T4 head assembly. *In vivo* characterization of the morphopoietic core. In: Dubow, M., editor. Bacteriophage Assembly. NY. USA: Alan R. Liss, Inc.; 1981. p. 127-137.
51. Showe MK, Black LW. Assembly core of bacteriophage T4: an intermediate in head formation. Nat. New Biol. 1973; 242(116):70–75. [PubMed: 4512235]
52. Laemmli UK, Beguin F, Gujer-Kellenberger G. A factor preventing the major head protein of bacteriophage T4 from random aggregation. J. Mol. Biol. 1970; 47(1):69–85. [PubMed: 5413343]
53. Laemmli UK, Molbert E, Showe M, Kellenberger E. Form-determining function of the genes required for the assembly of the head of bacteriophage T4. J. Mol. Biol. 1970; 49(1):99–113. [PubMed: 5450520]
54. Zeilstra-Ryalls J, Fayet O, Georgopoulos C. The universally conserved GroE (Hsp60) chaperonins. Annu. Rev. Microbiol. 1991; 45:301–325. [PubMed: 1683763]
55. Wunderli H, Couture E, Vince DA, Kellenberger E. Studies related to the head-maturation pathway of bacteriophages T4 And T2: II. Nuclear disruption, protein synthesis and particle formation with the mutant 43-.30-.46. J. Supramol. Struct. 1977; 7(2):163–190. [PubMed: 599934]
56. Schaerli C, Kellenberger. Head maturation pathway of bacteriophages T4 and T2. V. Maturable epsilon-particle accumulating an acridine-treated bacteriophage T4-infected cells. J. Virol. 1980; 33(2):830–844. [PubMed: 6997509]
57. Carrascosa JL, Kellenberger E. Head maturation pathway of bacteriophages T4 and T2. III. Isolation and characterization of particles produced by mutants in gene 17. J. Virol. 1978; 25(3): 831–844. [PubMed: 642076]
58. Laemmli UK, Johnson RA. Maturation of the head of bacteriophage T4. II. Head-related, aberrant tau-particles. J. Mol. Biol. 1973; 80(4):601–611. [PubMed: 4589646]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

59. Laemmli UK, Quittner SF. Maturation of the head of bacteriophage T4. IV. The proteins of the core of the tubular polyheads and *in vitro* cleavage of the head proteins. Virology. 1974; 62(2): 483–499. [PubMed: 4432376]
60. Castillo CJ, Hsiao CL, Coon P, Black LW. Identification and properties of bacteriophage T4 capsid-formation gene products. J. Mol. Biol. 1977; 110(3):585–601. [PubMed: 845961]
61. Showe MK, Isobe E, Onorato L. Bacteriophage T4 prehead proteinase. II. Its cleavage from the product of gene 21 and regulation in phage-infected cells. J. Mol. Biol. 1976; 107(1):55–69. [PubMed: 1003460]
62. Showe MK, Isobe E, Onorato L. Bacteriophage T4 prehead proteinase. I. Purification and properties of a bacteriophage enzyme which cleaves the capsid precursor proteins. J. Mol. Biol. 1976; 107(1):35–54. [PubMed: 12371]
63. Isobe T, Black LW, Tsugita A. Primary structure of bacteriophage T4 internal protein II and characterization of the cleavage upon phage maturation. J. Mol. Biol. 1976; 102(2):349–365. [PubMed: 1271467]
64. McNicol LA, Simon LD, Black LW. A mutation which bypasses the requirement for p24 in bacteriophage T4 capsid morphogenesis. J. Mol. Biol. 1977; 116(2):261–283. [PubMed: 599558]
65. Zhang Z, Kottadiel VI, Vafabakhsh R, et al. A promiscuous DNA packaging machine from bacteriophage T4. PLoS Biol. 2011; 9(2):e1000592. [PubMed: 21358801]
66. Rao VB, Black LW. Cloning, overexpression and purification of the terminase proteins gp16 and gp17 of bacteriophage T4: construction of a defined *in-vitro* DNA packaging system using purified terminase proteins. J. Mol. Biol. 1988; 200(3):475–488. [PubMed: 3294420]
67. Rao VB, Feiss M. The bacteriophage DNA packaging motor. Annu. Rev. Genet. 2008; 42:642–681.
68. Simpson AA, Tao Y, Leiman PG, et al. Structure of the bacteriophage f29 DNA packaging motor. Nature. 2000; 408(6813):745–750. [PubMed: 11130079]
69. Lebedev AA, Krause MH, Isidro AL, et al. Structural framework for DNA translocation via the viral portal protein. EMBO J. 2007; 26(7):1984–1994. [PubMed: 17363899]
70. Olia AS, Prevelige PE Jr, Johnson JE, Cingolani G. Three-dimensional structure of a viral genome-delivery portal vertex. Nat. Struct. Mol. Biol. 2011; 18(5):597–603. [PubMed: 21499245]
71. Donate LE, Herranz L, Secilla JP, Carazo JM, Fujisawa H, Carrascosa JL. Bacteriophage T3 connector: three-dimensional structure and comparison with other viral head-tail connecting regions. J. Mol. Biol. 1988; 201(1):91–100. [PubMed: 3262165]
72. Valpuesta JM, Fujisawa H, Marco S, Carazo JM, Carrascosa JL. Three-dimensional structure of T3 connector purified from overexpressing bacteria. J. Mol. Biol. 1992; 224(1):103–112. [PubMed: 1548694]
73. Sun S, Kondabagil K, Draper B, et al. The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. Cell. 2008; 135(7):1251–1262. [PubMed: 19109896]
74. Kuebler D, Rao VB. Functional analysis of the DNA-packaging/terminase protein gp17 from bacteriophage T4. J. Mol. Biol. 1998; 281(5):803–814. [PubMed: 9719636]
75. Rao VB, Mitchell MS. The N-terminal ATPase site in the large terminase protein gp17 is critically required for DNA packaging in bacteriophage T4. J. Mol. Biol. 2001; 314(3):401–411. [PubMed: 11846554]
76. Rossmann MG, Moras D, Olsen KW. Chemical and biological evolution of a nucleotide-binding protein. Nature. 1974; 250(463):194–199. [PubMed: 4368490]
77. Kanamaru S, Kondabagil K, Rossmann MG, Rao VB. The functional domains of bacteriophage T4 terminase. J. Biol. Chem. 2004; 279(39):40795–40801. [PubMed: 15265872]
78. Lin H, Simon MN, Black LW. Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging. J. Biol. Chem. 1997; 272(6):3495–3501. [PubMed: 9013596]
79. Lin H, Black LW. DNA requirements *in vivo* for phage T4 packaging. Virology. 1998; 242(1): 118–127. [PubMed: 9501053]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

80. Al-Zahrani AS, Kondabagil K, Gao S, Kelly N, Ghosh-Kumar M, Rao VB. The small terminase, gp16, of bacteriophage T4 is a regulator of the DNA packaging motor. J. Biol. Chem. 2009; 284(36):24490–24500. [PubMed: 19561086]
81. Catalano CE, Cue D, Feiss M. Virus DNA packaging: the strategy used by phage l. Mol. Microbiol. 1995; 16(6):1075–1086. [PubMed: 8577244]
82. Chai S, Lurz R, Alonso JC. The small subunit of the terminase enzyme of Bacillus subtilis bacteriophage SPP1 forms a specialized nucleoprotein complex with the packaging initiation region. J. Mol. Biol. 1995; 252(4):386–398. [PubMed: 7563059]
83. Takahashi H, Saito H. High-frequency transduction of pBR322 by cytosine-substituted T4 bacteriophage: evidence for encapsulation and transfer of head-to-tail plasmid concatemers. Plasmid. 1982; 8(1):29–35. [PubMed: 6755510]
84. Wilson GG, Young KY, Edlin GJ, Konigsberg W. High-frequency generalised transduction by bacteriophage T4. Nature. 1979; 280(5717):80–82. [PubMed: 15305587]
85. Büttner CR, Chechik M, Ortiz-Lombardia M, et al. Structural basis for DNA recognition and loading into a viral packaging motor. Proc. Natl Acad. Sci. USA. 2012; 109(3):811–816. [PubMed: 22207627]
86. Zhao H, Finch CJ, Sequeira RD, et al. Crystal structure of the DNA-recognition component of the bacterial virus Sf6 genome-packaging machine. Proc. Natl Acad. Sci. USA. 2010; 107(5):1971–1976. [PubMed: 20133842]
87. Sun S, Gao S, Kondabagil K, Xiang Y, Rossmann MG, Rao VB. Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. Proc. Natl Acad. Sci. USA. 2012; 109(3):817–822. [PubMed: 22207623]
88. Zhao H, Kamau YN, Christensen TE, Tang L. Structural and functional studies of the phage Sf6 terminase small subunit reveal a DNA-spooling device facilitated by structural plasticity. J. Mol. Biol. 2012; 423(3):413–426. [PubMed: 22858866]
89. Gao S, Rao VB. Specificity of interactions among the DNA-packaging machine components of T4-related bacteriophages. J. Biol. Chem. 2011; 286(5):3944–3956. [PubMed: 21127059]
90. Black LW, Rao VB. Structure, assembly, and DNA packaging of the bacteriophage T4 head. Adv. Virus Res. 2012; 82:119–153. [PubMed: 22420853]
91. Aebi U, Bijlenga R, van den Broek J, et al. The transformation of tau particles into T4 heads. II. Transformations of the surface lattice and related observations on form determination. J. Supramol. Struct. 1974; 2(2–4):253–275. [PubMed: 4612249]
92. Aebi UR, van Driel R, Bijlenga RKL, et al. Capsid fine structure of T-even bacteriophages. Binding and localization of two dispensable capsid proteins into the P23 surface lattice. J. Mol. Biol. 1977; 110(4):687–698. [PubMed: 859178]
93. Steven AC, Aebi U, Showe MK. Folding and capsomere morphology of the P23 surface shell of bacteriophage T4 polyheads from mutants in five different head genes. J. Mol. Biol. 1976; 102(3): 373–400. [PubMed: 775106]
94. Steven AC, Couture E, Aebi U, Showe MK. Structure of T4 polyheads. II. A pathway of polyhead transformation as a model for T4 capsid maturation. J. Mol. Biol. 1976; 106(1):187–221. [PubMed: 972397]
95. Mueller, JE.; Bryk, M.; Loizos, N.; Belfort, M. Homing endonucleases. In: Linn, SM.; Lloyd, RS.; Roberts, RJ., editors. The Nucleases. NY, USA: Cold Spring Harbor Press; 1993. p. 111-143.
96. Steven AC, Greenstone HL, Booy FP, Black LW, Ross PD. Conformational changes of a viral capsid protein. Thermodynamic rationale for proteolytic regulation of bacteriophage T4 capsid expansion, cooperativity, and super-stabilization by soc binding. J. Mol. Biol. 1992; 228(3):870–884. [PubMed: 1469720]
97. Ishii T, Yanagida M. Molecular organization of the shell of the T-even bacteriophage head. J. Mol. Biol. 1975; 97(4):655–660. [PubMed: 1185785]
98. Hendrix RW. Symmetry mismatch and DNA packaging in large bacteriophages. Proc. Natl Acad. Sci. USA. 1978; 75(10):4779–4783. [PubMed: 283391]
99. Baumann RG, Mullaney J, Black LW. Portal fusion protein constraints on function in DNA packaging of bacteriophage T4. Mol. Microbiol. 2006; 61(1):16–32. [PubMed: 16824092]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

100. Fuller DN, Raymer DM, Kottadiel V, Rao VB, Smith DE. Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. Proc. Natl Acad. Sci. USA. 2007; 104(43):16868–16873. [PubMed: 17942694]
101. Comolli LR, Spakowitz AJ, Siegerist CE, et al. Three-dimensional architecture of the bacteriophage phi29 packaged genome and elucidation of its packaging process. Virology. 2008; 371(2):267–277. [PubMed: 18001811]
102. Leiman PG, Arisaka F, van Raaij MJ, et al. Morphogenesis of the T4 tail and tail fibers. Virol. J. 2010; 7:355. [PubMed: 21129200]
103. Kikuchi Y, King J. Genetic control of bacteriophage T4 baseplate morphogenesis. I. Sequential assembly of the major precursor *in vivo* and *in vitro* . J. Mol. Biol. 1975; 99(4):645–672. [PubMed: 765481]
104. Kikuchi Y, King J. Genetic control of bacteriophage T4 baseplate morphogenesis. II. Mutants unable to form the central part of the baseplate. J. Mol. Biol. 1975; 99(4):673–694. [PubMed: 765482]
105. Kikuchi Y, King J. Genetic control of bacteriophage T4 baseplate morphogenesis. III. Formation of the central plug and overall assembly pathway. J. Mol. Biol. 1975; 99(4):695–716. [PubMed: 765483]
106. Plishker MF, Rangwala SH, Berget PB. Isolation of bacteriophage T4 baseplate proteins P7 and P8 and *in vitro* formation of the P10/P7/P8 assembly intermediate. J. Virol. 1988; 62(2):400–406. [PubMed: 3336066]
107. Ferguson PL, Coombs DH. Pulse-chase analysis of the *in vivo* assembly of the bacteriophage T4 tail. J. Mol. Biol. 2000; 297(1):99–117. [PubMed: 10704310]
108. Yap ML, Mio K, Leiman PG, Kanamaru S, Arisaka F. The baseplate wedges of bacteriophage T4 spontaneously assemble into hubless baseplate-like structure *in vitro*. J. Mol. Biol. 2010; 395(2): 349–360. [PubMed: 19896486]
109. Abuladze NK, Gingery M, Tsai J, Eiserling FA. Tail length determination in bacteriophage T4. Virology. 1994; 199(2):301–310. [PubMed: 8122363]
110. Kostyuchenko VA, Chipman PR, Leiman PG, Arisaka F, Mesyanzhinov VV, Rossmann MG. The tail structure of bacteriophage T4 and its mechanism of contraction. Nat. Struct. Mol. Biol. 2005; 12(9):810–813. [PubMed: 16116440]
111. Kanamaru S, Leiman PG, Kostyuchenko VA, et al. Structure of the cell-puncturing device of bacteriophage T4. Nature. 2002; 415(6871):553–557. [PubMed: 11823865]
112. Kostyuchenko VA, Navruzbekov GA, Kurochkina LP, Strelkov SV, Mesyanzhinov VV, Rossmann MG. The structure of bacteriophage T4 gene product 9: the trigger for tail contraction. Structure. 1999; 7(10):1213–1222. [PubMed: 10545330]
113. Leiman PG, Shneider MM, Mesyanzhinov VV, Rossmann MG. Evolution of bacteriophage tails: structure of T4 gene product 10. J. Mol. Biol. 2006; 358(3):912–921. [PubMed: 16554069]
114. Leiman PG, Kostyuchenko VA, Shneider MM, Kurochkina LP, Mesyanzhinov VV, Rossmann MG. Structure of bacteriophage T4 gene product 11, the interface between the baseplate and short tail fibers. J. Mol. Biol. 2000; 301(4):975–985. [PubMed: 10966799]
115. Van Raaij MJ, Schoehn G, Burda MR, Miller S. Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. J. Mol. Biol. 2001; 314(5):1137–1146. [PubMed: 11743729]
116. Thomassen E, Gielen G, Schutz M, et al. The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. J. Mol. Biol. 2003; 331(2):361–373. [PubMed: 12888344]
117. Leiman PG, Shneider MM, Kostyuchenko VA, Chipman PR, Mesyanzhinov VV, Rossmann MG. Structure and location of gene product 8 in the bacteriophage T4 baseplate. J. Mol. Biol. 2003; 328(4):821–833. [PubMed: 12729757]
118. Aksyuk AA, Leiman PG, Shneider MM, Mesyanzhinov VV, Rossmann MG. The structure of gene product 6 of bacteriophage T4, the hinge-pin of the baseplate. Structure. 2009; 17(6):800–808. [PubMed: 19523898]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

119. Blake CCF, Koenig DF, Mair GA, North ACT, Phillips DC, Sarma VR. Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Å resolution. Nature. 1965; 206(4986):757–761. [PubMed: 5891407]
120. Kanamaru S, Gassner NC, Ye N, Takeda S, Arisaka F. The C-terminal fragment of the precursor tail lysozyme of bacteriophage T4 stays as a structural component of the baseplate after cleavage. J. Bacteriol. 1999; 181:2739–2744. [PubMed: 10217762]
121. Kanamaru S, Ishiwata Y, Suzuki T, Rossmann MG, Arisaka F. Control of bacteriophage T4 tail lysozyme activity during the infection process. J. Mol. Biol. 2005; 346:1013–1020. [PubMed: 15701513]
122. Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. Nature. 2013; 500(7462):350–353. [PubMed: 23925114]
123. Aksyuk AA, Leiman PG, Kurochkina LP, et al. The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. EMBO J. 2009; 28(7):821–829. [PubMed: 19229296]
124. Fokine A, Zhang Z, Kanamaru S, et al. The molecular architecture of the bacteriophage T4 neck. J. Mol. Biol. 2013; 425(10):1731–1744. [PubMed: 23434847]
125. Hashemolhosseini S, Stierhof YD, Hindennach I, Henning U. Characterization of the helper proteins for the assembly of tail fibers of coliphages T4 and l. J. Bacteriol. 1996; 178(21):6258–6265. [PubMed: 8892827]
126. King J, Wood WB. Assembly of bacteriophage T4 tail fibers: the sequence of gene product interaction. J. Mol. Biol. 1969; 39(3):583–601. [PubMed: 5390559]
127. Cerritelli ME, Wall JS, Simon MN, Conway JF, Steven AC. Stoichiometry and domainal organization of the long tail-fiber of bacteriophage T4: a hinged viral adhesin. J. Mol. Biol. 1996; 260(5):767–780. [PubMed: 8709154]
128. Bartual SG, Otero JM, Garcia-Doval C, et al. Structure of the bacteriophage T4 long tail fiber receptor-binding tip. Proc. Natl Acad. Sci. USA. 2010; 107(47):20287–20292. [PubMed: 21041684]
129. Montag D, Hashemolhosseini S, Henning U. Receptor-recognizing proteins of T-even type bacteriophages. The receptor-recognizing area of proteins 37 of phages T4 TuIa and TuIb. J. Mol. Biol. 1990; 216(2):327–334. [PubMed: 2147721]
130. Coombs DH, Eiserling FA. Studies on the structure, protein composition and assembly of the neck of bacteriophage T4. J. Mol. Biol. 1977; 116(3):375–405. [PubMed: 592389]
131. Tao Y, Strelkov SV, Mesyanzhinov VV, Rossmann MG. Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain. Structure. 1997; 5(6):789–798. [PubMed: 9261070]
132. Strelkov SV, Tao Y, Shneider MM, Mesyanzhinov VV, Rossmann MG. Structure of bacteriophage T4 fibritin M: a troublesome packing arrangement. Acta Crystallogr. D Biol. Crystallogr. 1998; 54(Pt 5):805–816. [PubMed: 9757094]
133. Boudko SP, Strelkov SV, Engel J, Stetefeld J. Design and crystal structure of bacteriophage T4 mini-fibritin NCCF. J. Mol. Biol. 2004; 339(4):927–935. [PubMed: 15165860]
134. Efimov VP, Nepluev IV, Sobolev BN, et al. Fibritin encoded by bacteriophage T4 gene wac has a parallel triple-stranded a-helical coiled-coil structure. J. Mol. Biol. 1994; 242(4):470–486. [PubMed: 7932704]
135. Letarov AV, Londer YY, Boudko SP, Mesyanzhinov VV. The carboxy-terminal domain initiates trimerization of bacteriophage T4 fibritin. Biochemistry (Mosc.). 1999; 64(7):817–823. [PubMed: 10424907]
136. Conley MP, Wood WB. Bacteriophage T4 whiskers: a rudimentary environment-sensing device. Proc. Natl Acad. Sci. USA. 1975; 72(9):3701–3705. [PubMed: 242007]
137. Sun S, Kondabagil K, Gentz PM, Rossmann MG, Rao VB. The structure of the ATPase that powers DNA packaging into bacteriophage T4 procapsids. Mol. Cell. 2007; 25(6):943–949. [PubMed: 17386269]
138. Wikoff WR, Liljas L, Duda RL, Tsuruta H, Hendrix RW, Johnson JE. Topologically linked protein rings in the bacteriophage HK97 capsid. Science. 2000; 289(5487):2129–2133. [PubMed: 11000116]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

139. Morais MC, Kanamaru S, Badasso MO, et al. Bacteriophage f29 scaffolding protein gp7 before and after prohead assembly. Nat. Struct. Biol. 2003; 10:572–576. [PubMed: 12778115]

140. Trus BL, Cheng N, Newcomb WW, Homa FL, Brown JC, Steven AC. Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. J. Virol. 2004; 78(22): 12668–12671. [PubMed: 15507654]

141. Fokine A, Bowman VD, Battisti AJ, et al. Cryo-electron microscopy study of bacteriophage T4 displaying anthrax toxin proteins. Virology. 2007; 367(2):422–427. [PubMed: 17624389]

142. Ren ZJ, Lewis GK, Wingfield PT, Locke EG, Steven AC, Black LW. Phage display of intact domains at high copy number: a system based on SOC, the small outer capsid protein of bacteriophage T4. Protein Sci. 1996; 5(9):1833–1843. [PubMed: 8880907]

143. Ren ZJ, Tian CJ, Zhu QS, et al. Orally delivered foot-and-mouth disease virus capsid protomer vaccine displayed on T4 bacteriophage surface: 100% protection from potency challenge in mice. Vaccine. 2008; 26(11):1471–1481. [PubMed: 18289743]

144. Jiang J, Abu-Shilbayeh L, Rao VB. Display of a PorA peptide from *Neisseria meningitidis* on the bacteriophage T4 capsid surface. Infect. Immun. 1997; 65(11):4770–4777. [PubMed: 9353063]

145. Sathaliyawala T, Rao M, Maclean DM, Birx DL, Alving CR, Rao VB. Assembly of human immunodeficiency virus (HIV) antigens on bacteriophage T4: a novel *in vitro* approach to construct multicomponent HIV vaccines. J. Virol. 2006; 80(15):7688–7698. [PubMed: 16840347]

146. Wu J, Tu C, Yu X, et al. Bacteriophage T4 nanoparticle capsid surface SOC and HOC bipartite display with enhanced classical swine fever virus immunogenicity: a powerful immunological approach. J. Virol. Methods. 2007; 139(1):50–60. [PubMed: 17081627]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

EXECUTIVE SUMMARY

T4 head structure

- The 1150 Å-long and 850 Å-wide head consists of 155 hexameric capsid protein gp23*, 11 pentameric vertex protein gp24*, a dodecameric portal protein gp20, 155 Hoc and 870 Soc proteins.
- The large terminase gp17 has an N-terminal ATPase domain for ATP hydrolysis and DNA translocation, and a C-terminal translocation and nuclease domain.
- An electrostatic force-driven mechanism was proposed for DNA packaging. Alternating tensed and relaxed conformational states generated through charge attraction between the ATPase and nuclease domains of gp17 have been proposed to pump the 172 kbp DNA into the head.

T4 tail structure

- The tail tube protein gp19 polymerizes on the baseplate (520 Å-diameter, 270 Å-high) to form the tail tube, which is finally capped by the tail-tube terminator protein gp3 hexamer. The gp18 polymerizes around the tail tube to form the contractile sheath. The sheath assembly is finalized by attachment of a gp15 hexamer to the last row of gp18 and gp3.
- The cell-puncturing device consists of trimeric gp5-gp27 and monomeric gp5.4. It functions in the manner of a drill to pierce the outer cell membrane followed by lysis of the peptidoglycan layer located in the host's periplasm.
- During infection, the baseplate transforms from a dome shape structure to a star shape structure. The tail sheath contracts. The tail tube rotates by almost one complete turn to drill through the host cell membrane and the genome packaged inside the head passes through the tube into cytosol of the host cell.

Tail fiber structure, collar & whisker

- Six 1450 Å-long long tail fibers, each consisting of gp34, gp35, gp36 and gp37, are bound to the baseplate. The C-terminal region of gp37 forms the receptor-recognizing tip to each long tail fiber.
- Twelve fibritin molecules attach to the phage neck to form the collar and whiskers. Fibritin molecules are sensors to recognize optimal conditions for infection.

Conclusion

- Recent structural analyses provide more complete knowledge of T4 morphogenesis and elucidate some of the molecular mechanisms involved in the T4 life cycle.
- Conserved protein folds related to specific biological functions imply co-evolution among different biological systems.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

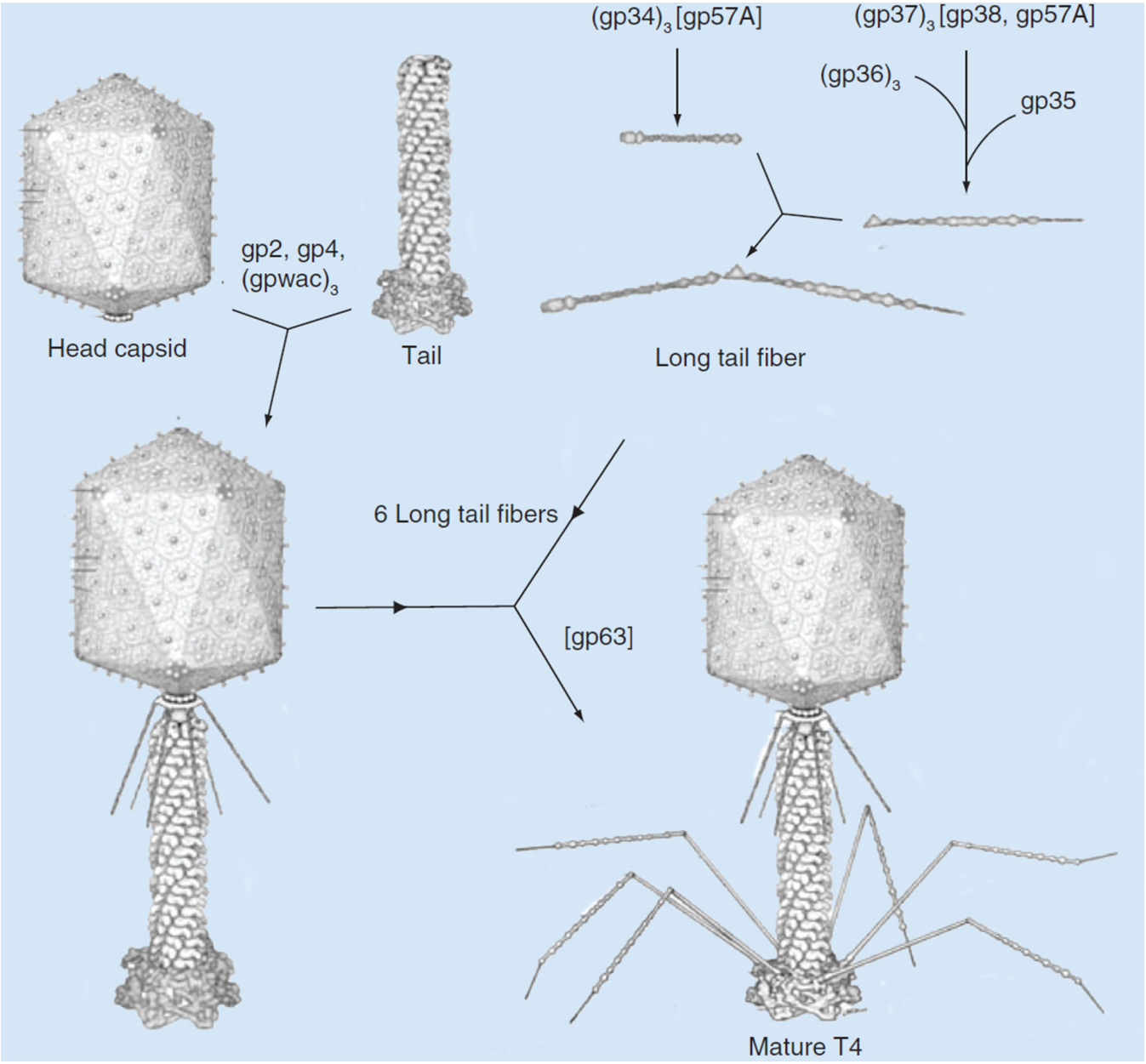

Figure 1. Assembly pathway of bacteriophage T4

The assembly of T4 can be divided into three independent subassemblies: the head, the tail and the long tail fibers. The tail binds to the head followed by attachment of the fibritin protein at the neck region. Six long tail fibers then attach to form a viable T4 virion. Adapted with permission from Figure 5 in [23]. © Cellular and Molecular Life Sciences (2003), reprinted with permission from Springer Science and Business Media.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

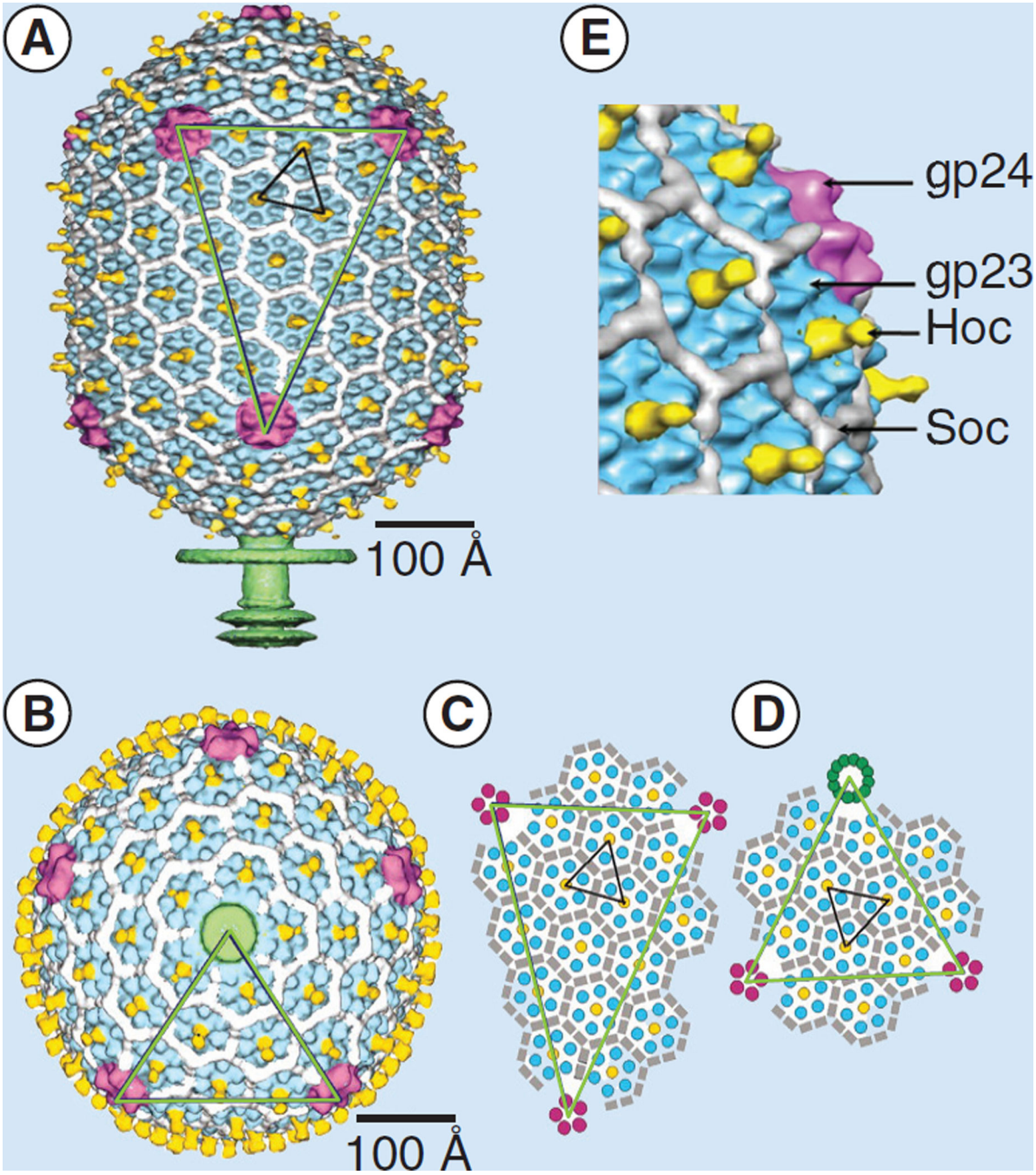

Figure 2. The structure of the bacteriophage T4 head

The symmetry of the gp23* major capsid protein shell is characterized by triangulation numbers T_{end} = 13 *laevo* and T_{mid} = 20 (h_1 = 3, k_1 = 1, h_2 = 4 and k_2 = 2) [26,27]. The facet triangles are shown in green and the basic triangles are shown in black. **(A)** Shaded surface representation of the cryo-electron microscopy reconstruction viewed perpendicular to the fivefold axis. gp23* is shown in blue, gp24* is in magenta, Soc is in white, Hoc is in yellow and the tail is in green. **(B)** View of the reconstruction along the fivefold axis with the portal vertex toward the observer; the tail has been cut away at the level shown by the black arrow

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

in **(A)**. Proteins are colored as described for **(A)**. **(C)** Schematic representation of the distribution of proteins in the elongated midsection facet. **(D)** Schematic representation of an end-cap facet. Proteins are colored as described for **(A)** except the Soc molecules are shown as gray rectangles. **(E)** A closer look of the distribution of proteins on the head.
Adapted with permission from [22]. © National Academy of Sciences, USA (2004).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

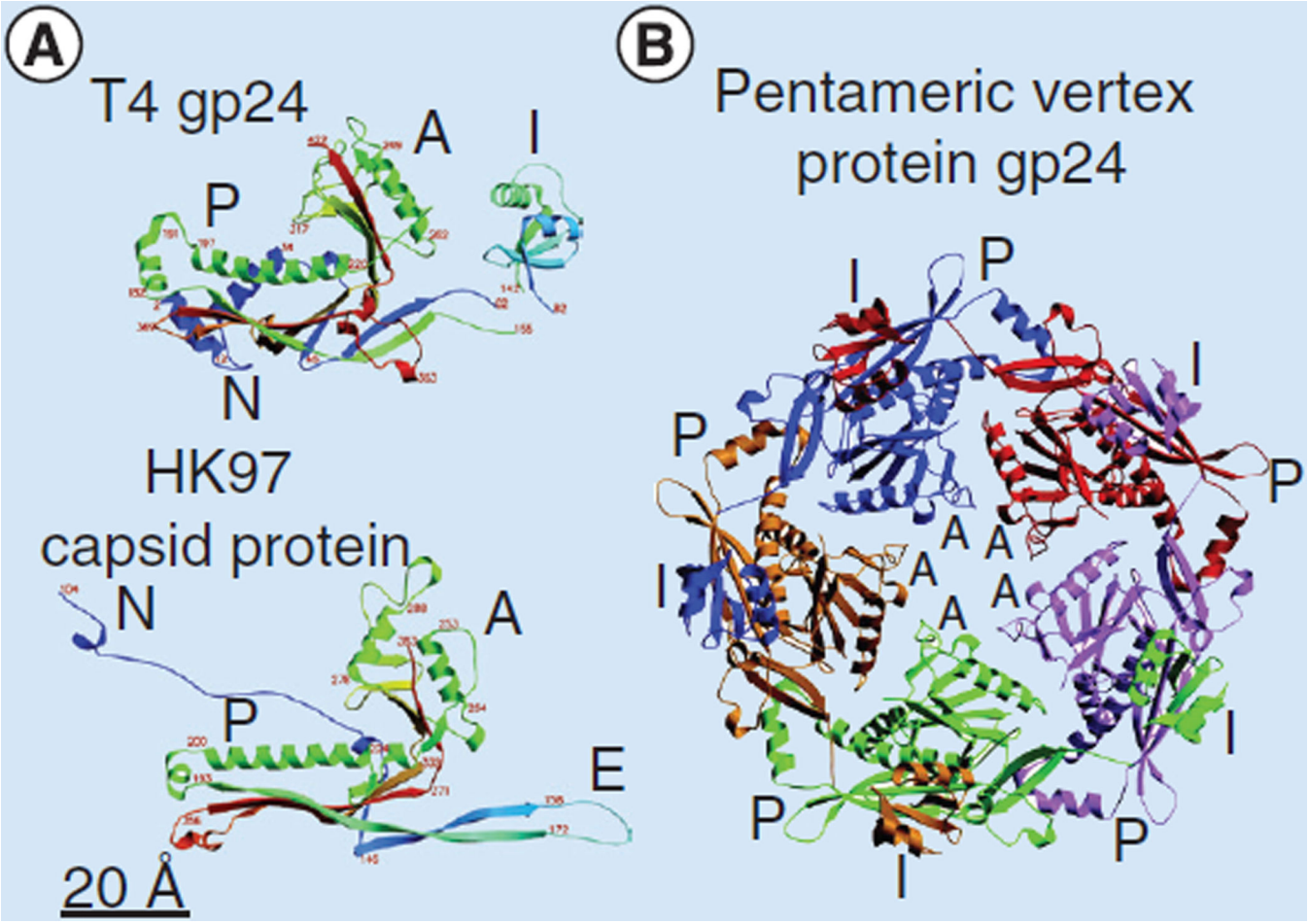

Figure 3. Crystal structure of the vertex protein gp24

(A) Ribbon diagrams show the structure of gp24 (upper) and HK97 capsid protein (lower). Colors range from blue at the amino end to red at the carboxyl end of the polypeptides. The P, A, and I domains, as well as the E and the N, are indicated. The linkers between the I and P domains of gp24 are disordered in the crystal structure. **(B)** Ribbon drawing of a gp24* pentamer determined by superimposing the gp24* crystal structure onto an HK97 pentamer. Each of the five molecules is shown in a different color. The P, A and I domains are shown. A: Axial; E: E-loop; I: Insertion; N: N-terminal domain; P: Peripheral.

Reproduced with permission from [31]. © National Academy of Sciences, USA (2005).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

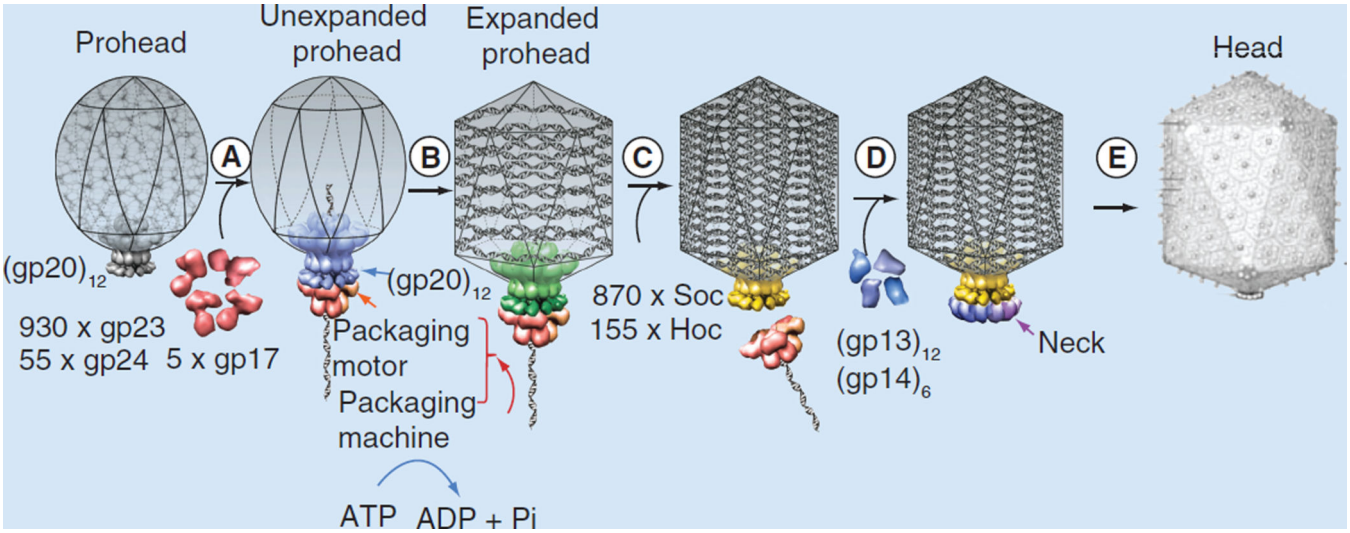

Figure 4. T4 head assembly

The dodecameric portal vertex acts as an initiator of head assembly. The major capsid protein assembles around a scaffolding core into a prohead. The core is removed by proteolysis to produce an empty unexpanded prohead. **(A)** The packaging machine-DNA complex docks on portal and initiates packaging. **(B)** The prohead expands after approximately 10–25% of the DNA is packaged. **(C)** When the head is full, the packaging machine cuts the DNA and dissociates from the head. **(D)** The neck proteins, gp13 and gp14, assemble on the portal to seal the DNA-full head. **(E)** The accessory proteins, Hoc and Soc, bind to the head. The head is ready to be joined to the tail.

Adapted with permission from Springer Science and Business Media [65] and Figure 5 in [23]. © Cellular and Molecular Life Sciences (2003).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

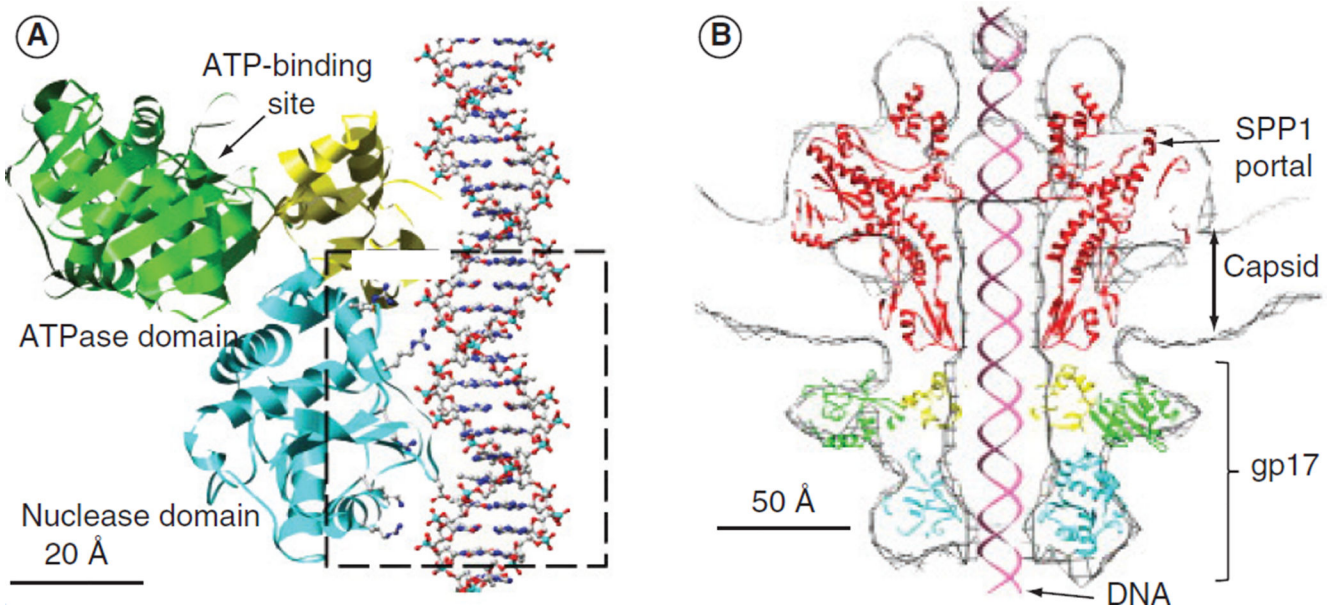

Figure 5. Structure of the phage T4 DNA-packaging machine

(A) A model of B-form polyA-polyT DNA molecule is shown bound to the T4 gp17 large terminase in the packaging mode. The N-terminal subdomain I is shown in green, the subdomain II is shown in yellow and the C-terminal domain is shown in cyan. The dashed box indicates potential interactions between the gp17 molecule and the DNA. **(B)** Structure of the dodecameric SPP1 portal (red) (based on the crystal structure of the 13-mer) and crystal structure of T4 gp17 were fitted into the T4 procapsid-gp17 complex cryo-electron microscopy density.

Adapted with permission from Elsevier [73]. © *Cell* (2008).

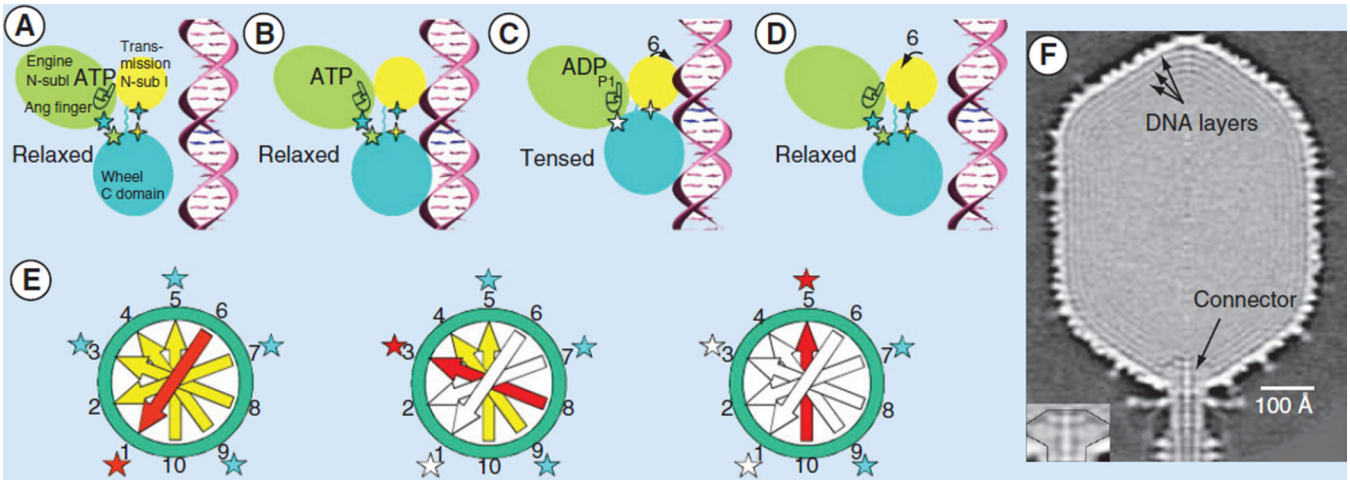

Figure 6. The DNA-packaging mechanism

Panels **(A–D)** relate to the sequence of events that occur in a single gp17 molecule. The gp17 N-terminal subdomain I, subdomain II and C-terminal domain are represented as green, yellow and cyan ovals, respectively. The five-pointed stars show the charge interactions between the N-terminal subdomain I and the C-terminal domain. The four-pointed stars show the charge interaction between the N-terminal subdomain II and the C-terminal domain. The flexible linker between N- and C-terminal domains is represented by a wiggly cyan line. **(A)** The gp17 C-terminal domain is ready to bind DNA. **(B)** The C-terminal domain, when bound to the DNA, brings the DNA closer to the N-terminal domain of the same subunit. Conformational change in the N-terminal domain causes Arg162 to be placed into the ATPase active center in preparation for hydrolysis. **(C)** Hydrolysis of ATP has rotated the N-terminal subdomain II by approximately 6°, thereby aligning the charge pairs resulting in an electrostatic attraction that moves the C-terminal domain and the DNA 6.8 Å (equivalent to the distance between two basepairs) closer to the N-terminal domain and into the capsid. **(D)** ADP and Pi are released and the C-terminal domain returns to its original position. DNA is released and is aligned to bind the C-terminal domain of the neighboring subunit. **(E)** This panel relates to the synchronization of the five gp17 molecules located around the special vertex of the procapsid. Successive DNA base pairs are indicated by yellow arrows outside the procapsid, red entering the procapsid, and white inside the procapsid. The surrounding five gp17 molecules are shown as stars. The red star represents gp17-ATPase-hydrolyzing ATP, the blue star represents ATPase that is ready to hydrolyze ATP and the white star represents ATPase that has already hydrolyzed ATP. Left: Hydrolysis of ATP at position 1 translocating two base pairs into the procapsid. Middle: Hydrolysis of ATP at position 3 causing the translocation of further two base pairs into the procapsid. Right: The ATP at the position 5 is ready to be hydrolyzed. **(F)** The central section of the reconstructed cryo-electron microscopy density map viewed from a direction, perpendicular to the fivefold axis. The concentric layers beneath the outer capsid shell are attributed to the densely packaged dsDNA. The spacing between successive layers is approximately 25 Å. The high-density region in the capsid interior next to the portal vertex is attributed to the gp20 connector and is enlarged in the Inset.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(A–E) Adapted with permission from Elsevier [73], © *Cell* (2008); and **(F)** adapted with permission from [22], © National Academy of Sciences, USA (2004).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

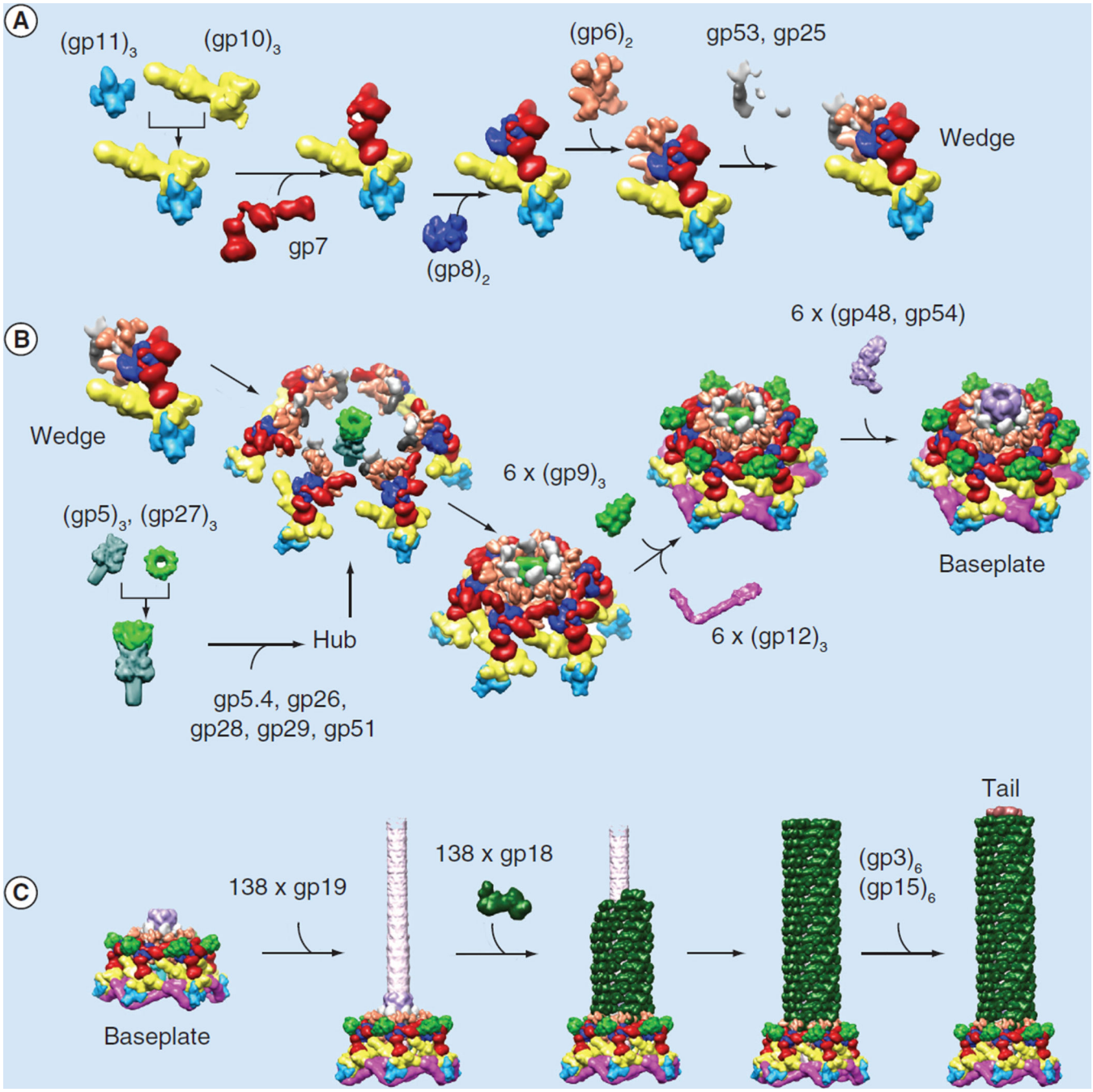

Figure 7. Assembly of the tail

(A) The assembly of a wedge. **(B)** Six wedges surround the central hub to form a baseplate. **(C)** A baseplate serves as the platform for polymerization of tail tube and tail sheath. A tail is completed after capping of the tail terminator proteins.

Adapted with permission from [102].

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

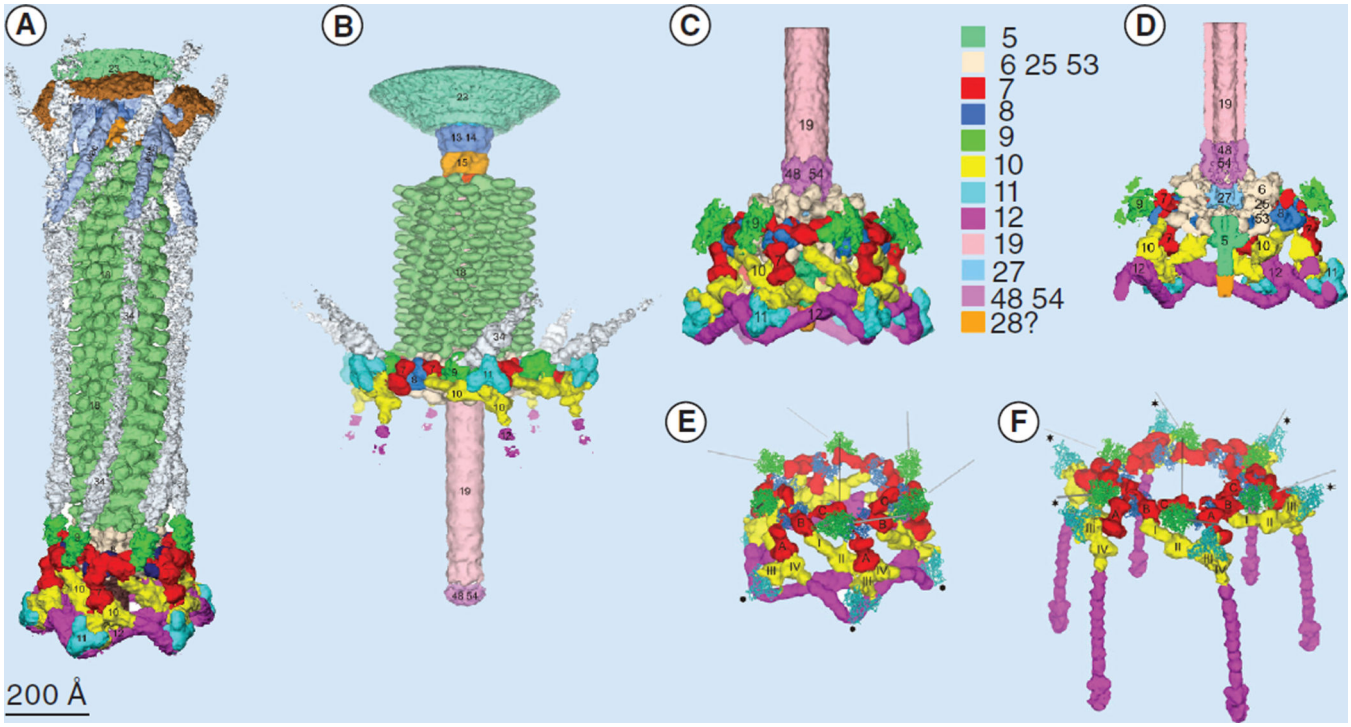

Figure 8. CryoEM reconstructions of the T4 tail and baseplate

The tail in an extended **(A)** and a contracted **(B)** conformation. **(C)** The baseplate and proximal part of the tail tube. Colors identity proteins labeled with their respective gene numbers, as shown on the right-hand side. **(D)** Cross-section view of the baseplate. Structure of the baseplate periphery of the baseplate in the dome shape **(E)** and the star shape **(F)** conformations.

(A) Adapted with permission from [110] © Nature Publishing Group (2005); **(B, E & F)** adapted with permission from Elsevier [21], © (2004); **(C & D)** adapted with permission from [20] © Nature Publishing Group.

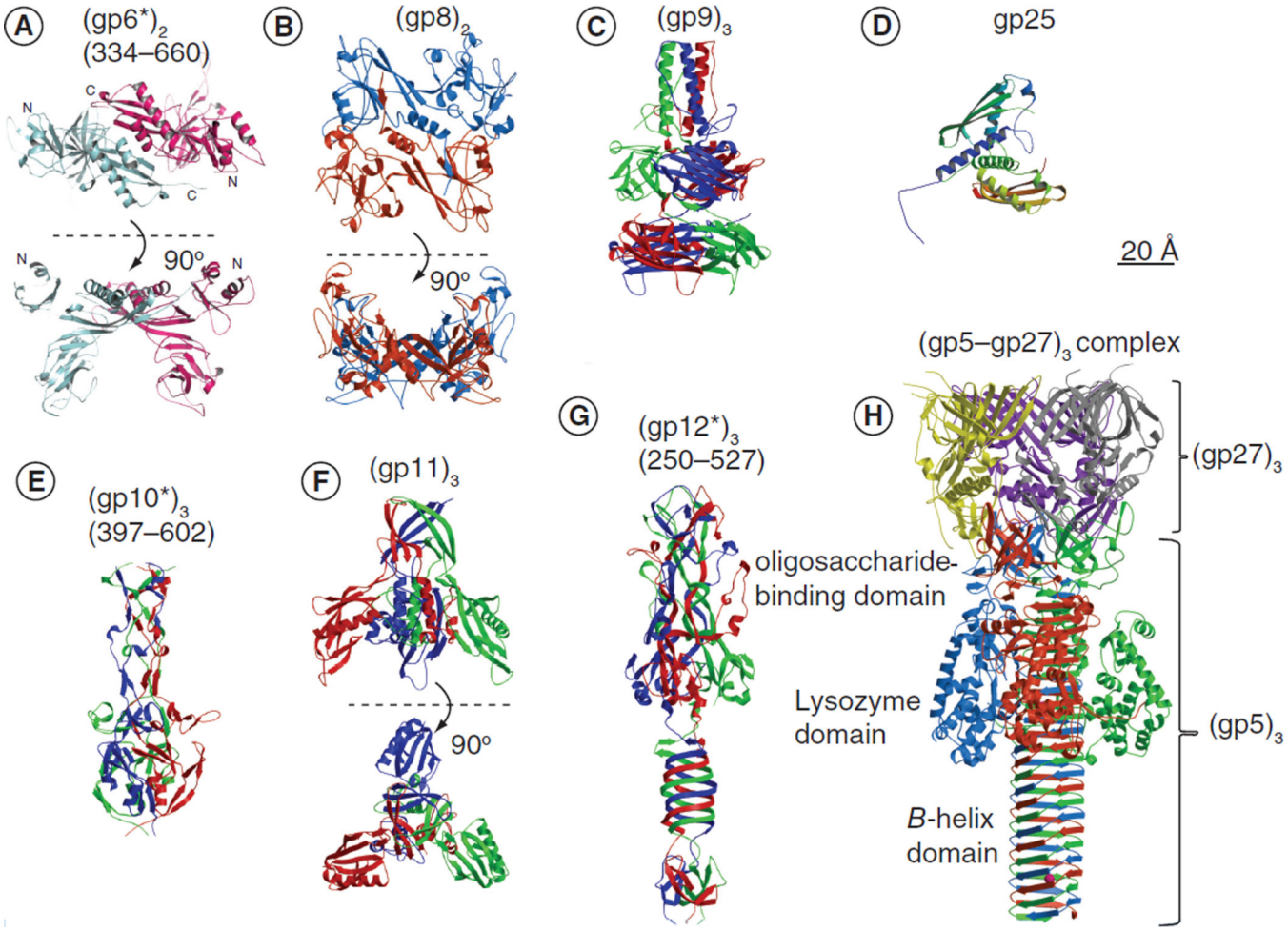

Figure 9. Crystal structures of the tail proteins

(A, B, D & E) are the wedge proteins. **(C)** is the adaptor of long tail fiber. **(F)** is the adaptor of short tail fiber. **(G)** is the short tail fiber. **(H)** is the cell-puncturing device. The asterisk (*) after the protein name denotes that the crystal structure is available for the C-terminal fragment of the protein. Residue numbers comprising the solved structure are given in parentheses.

(A–C, E–G) Adapted with permission from [102]; and **(H)** adapted with permission from [111], © Nature Publishing Group (2002).

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

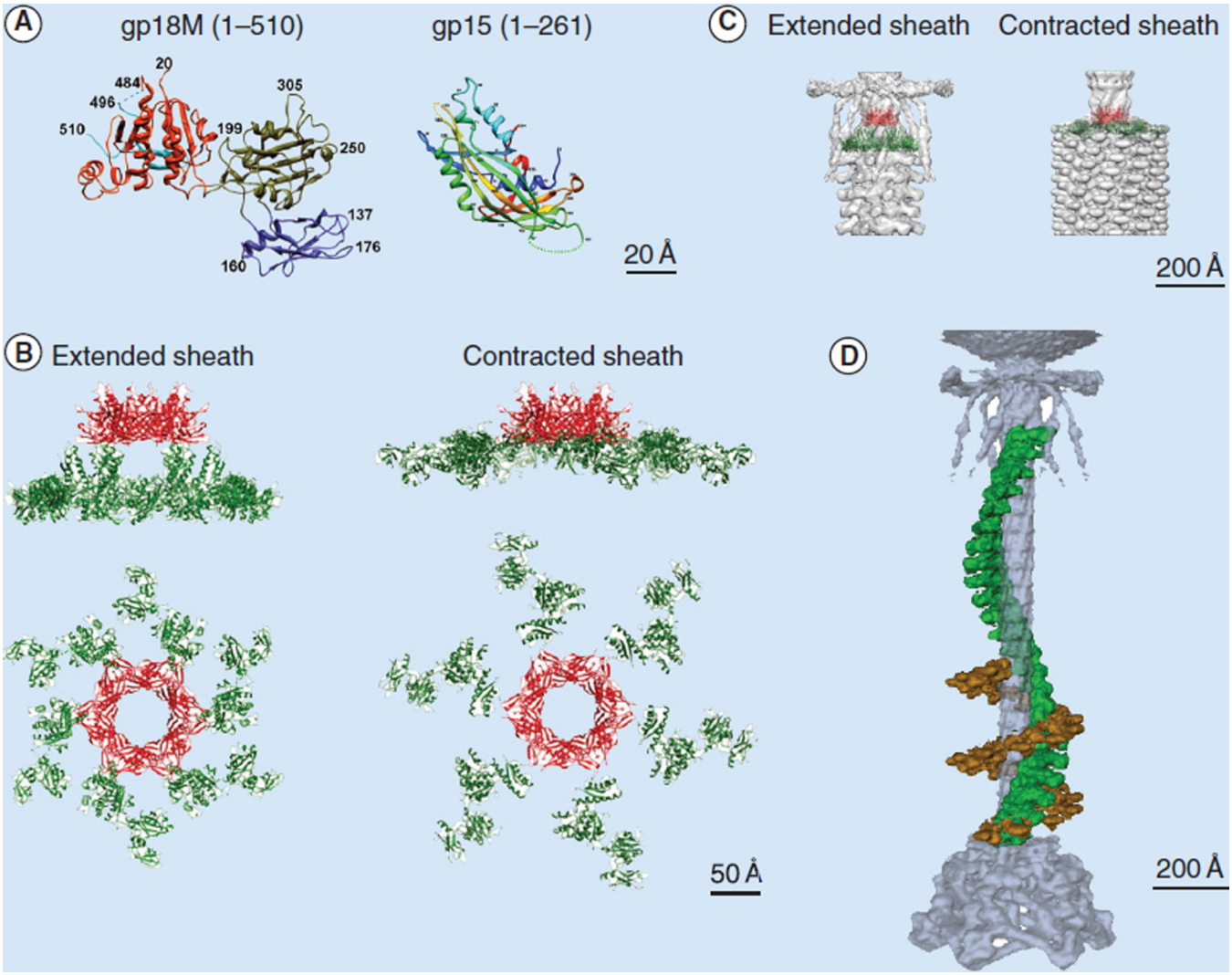

Figure 10. Crystal structures of gp18 and gp15

(A) Ribbon diagrams of gp18M and gp15 monomers. The three domains of gp18M are shown in blue (domain I), olive green (domain II) and orange red (domain III); the β-hairpin (residues 454–470) and the last 14 C-terminal residues are shown in cyan. Gp15 represented in rainbow color running from N terminus (blue) to C terminus (red). The broken lines indicate protein regions that are disordered in the crystals. **(B)** Relative positions of the gp15 and gp18 molecules in the extended and contracted T4 tails, viewed from the side (upper) and from the top (lower) of the phage. A model of the entire gp18 molecule was created based on the crystal structure of gp18M and the prophage tail sheath protein LIN1278. The models of gp18 molecules belonging to the topmost ring of the contractile sheath are shown in green. The gp15 hexamer is shown in red. **(C)** The gp15 and gp18 molecules are fitted into the cryo-electron microscopy reconstructions of the extended and contracted tails. **(D)** A helical strand of gp18 in the extended (green) and contracted (brown) tail. The hexagonal baseplate, tail tube, whiskers and collar are shown in gray-blue.

The left-hand side of **(A)** adapted with permission from John Wiley and Sons [123], © European Molecular Biology Organization (2009); **(A [right-hand side], B & C)** adapted

with permission from Elsevier [124], © (2013); and **(D)** adapted with permission from [110], © Nature Publishing Group (2005).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

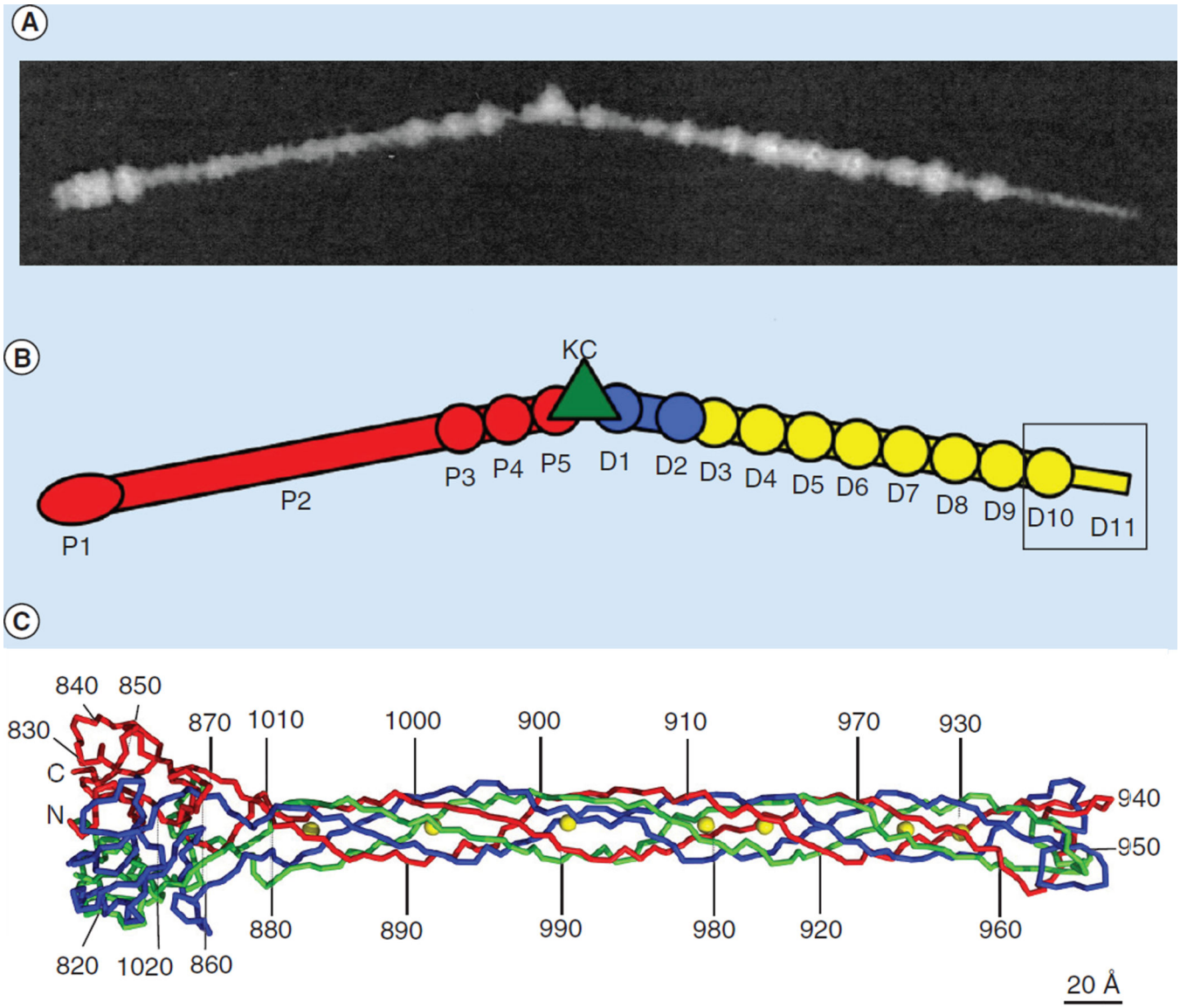

Figure 11. T4 long tail fiber

(A) Image of a long tail fiber. **(B)** Domain organization of the long tail fiber. Domains of the proximal tail fiber are named P1–5 and of the distal half D1–11; gp35, or the KC is represented as a green triangle. Crystal structure of D10 and D11 (box) has been determined. **(C)** Structure of gp37 (residues 785–1026). Three chains in the trimeric protein are colored red, green and blue, respectively, whereas iron ions are shown in yellow. The N- and C-termini and every 10th residue of chain A are labeled.

KC: Knee cap.

(A) Adapted with permission from Elsevier [127], © Elsevier (1996); **(B)** adapted with permission from [102] and **(C)** adapted with permission from National Academy of Sciences, USA [128].

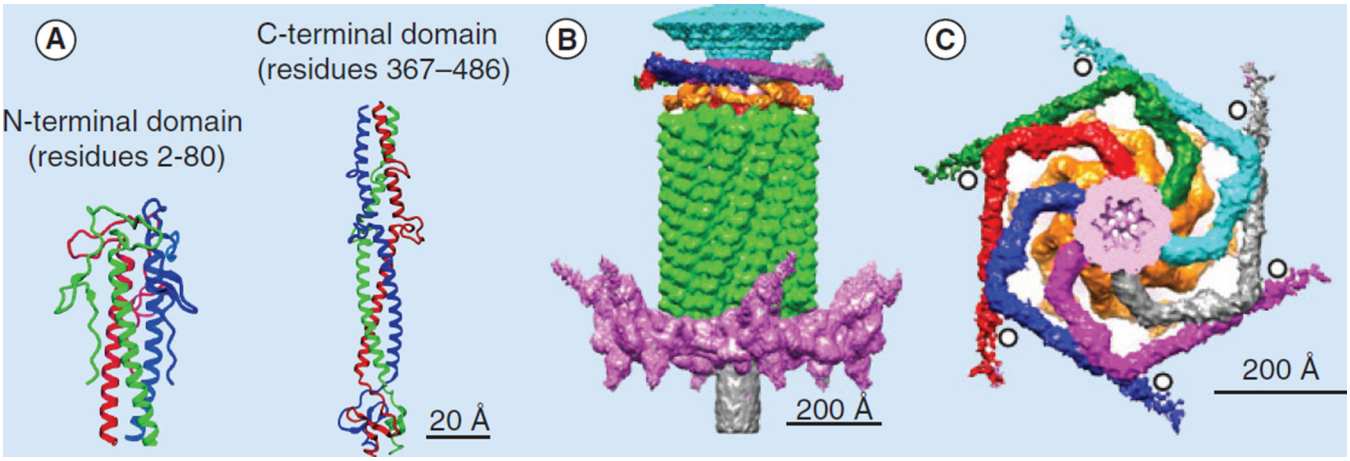

Figure 12. Structure of the collar and whiskers

(A) Crystal structure of N-terminal and C-terminal domains of the fibritin molecule. **(B)** Side view of the cryo-electron microscopy reconstruction of the contracted T4 tail. The baseplate, tail sheath and tail tube are colored magenta, green and gray, respectively. The phage whiskers are colored gold. Only parts of the whiskers are visible in the reconstruction because of their flexibility. **(C)** Top view of the collar and whisker structure. Surfaces of six different fibritin molecules forming the phage collar are colored green, cyan, gray, magenta, blue and red. The black circles show the positions where the retracted long tail fibers interact with the collar (in virus with extended tails).

(A) Adapted with permission from Elsevier [133] © Elsevier (2004); and **(B & C)** adapted with permission from Elsevier [124] © Elsevier (2013).

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript